Imaging Mass Spectrometry of the *Drosophila melanogaster* Wing Imaginal Disc

---

Dissertation
zur
Erlangung der naturwissenschaftlichen Doktorwürde
(Dr. sc. nat.)
vorgelegt der
Mathematisch-naturwissenschaftlichen Fakultät
der
Universität Zürich
von

Florian Marty

von

Gampel-Bratsch VS

Promotionskomitee
Prof. Dr. Konrad Basler
Prof. Dr. Ron M.A. Heeren
Prof. Dr. Reinhard Furrer
Dr. Markus Stoeckli
Dr. Erich Brunner (Leiter der Dissertation)

Zürich, 2014

Dedicated to my family Erich, Miranda, Katja, and Patrick and to Dr. Erich Brunner.

- It always seems impossible until it's done -

Nelson Mandela (1918 – 2013)

Table of Contents

Table of Figures

Abbreviations

A/P	Anterior/posterior
B/W	Body-wall/wing-blade
BG	Background
CARS	Coherent anti-Stokes Raman Spectroscopy
CCA	Canonical cross correlation analysis
CV	Canonical covariates
D/V	Dorsal/ventral
DESI	Desorption/Electrospray ionization
DG/DAG	Diacylglycerol
DGK	Diacylglycerolkinase
dh	Dorsal hinge
dpp	Decapentaplegic
ECM	Extracellular matrix
EGFR	Epidermal Growth Factor Receptor
en	Engrailed
ER	Endoplastamtic reticulum
ESI	Electrospray ionization
FT	Fourier Transform
GLM	Generalized linear model
hh	Hedgehog
ICR	Ion Cyclotron Resonance
IP_3	Inositol-3,4,5-triphosphate
Iro-C	Iroquois Complex
LC	Liquid chromatography
LDA	Linear discriminant analysis
LMIG	Liquid metal ion gun
Ln	Lanthanides
MALDI	Matrix-assisted laser desorption/ionization
MCP	Micro channel plate detector
MetA-SIMS	Metal-assisted SIMS
MG	Monoacylglycerol
msh or drp	Drop
MSI	Imaging mass spectrometry, Mass spectrometric imaging
Nd	Neodymium
NNMF	Non-negative matrix-factorization
PC	Principal component
PCA	Principal component analysis
PDMS	Polydimethylsiloxan
ROI	Region of interest

RS	Region of sample
SIMS	Secondary ion mass spectrometry
TG/TAG	Triacylglycerol
ToF	Time of flight
wg	Wingless

Summary

The development of an entire organism from a single fertilized egg cell is a fascinating process. To generate a complex organism such as the fruit fly *Drosophila melanogaster* tight control mechanisms are needed. Establishing the overall body plan around major axes helps to govern and control development. The organism, and by extension all major organs are organized by the same coordinate system: anterior-posterior (A/P) and dorsal-ventral (D/V). In *Drosophila melanogaster* this phenomenon has been extensively studied using the wing primordia as a model system. The so-called wing imaginal disc is patterned along the anterior-posterior axis by restricted activity of the selector gene *engrailed* (en), which is expressed only in posterior cells. Along the dorsal-ventral axis dorsal identity is specified by the selector gene *apterous,* which is restricted to dorsal cells. In addition to these classical axes, the wing imaginal disc is further subdivided by non-classical boundaries. These boundaries are often visible as morphological features such as tissue folds. A prominent example is the division between future body wall and wing blade (B/W). Although the mechanism of compartment boundary formation has been extensively studied genetically, much remains to be discovered. In particular, the role of small molecules has not yet been investigated.

In this work, we analyzed third instar wing imaginal discs with time of flight secondary ion imaging mass spectrometry (ToF-SIMS-MSI) to reveal the spatial distribution of small molecules on the tissue. Their co-localization to classical and non-classical compartments was verified with overlay images of green fluorescence-labeled (GFP) posterior compartment or co-registration with morphological features. Coherent anti-stokes Raman spectroscopy was applied to define the class of molecules observed. Unambiguous identification of molecules with distinctive spatial distributions was achieved generating a lipid catalogue of the wing imaginal disc. Comparison of the ToF-SIMS data with this catalogue and published literature lead to the identification of several lipids. These may have a function in formation and maintenance of the compartment boundaries in the wing imaginal disc. Additionally, two methods were developed to facilitate the interpretation of distorted imaging mass spectrometry data sets. Multiple collaborations were set up to perform the experiments herein:

- ToF-SIMS-MSI in positive and negative ion mode showed differential distribution of small molecules on the wing imaginal disc
- Developing two image correction algorithms improved interpretation of partially distorted data sets
- Co-localization of principle components (PCs) with GFP marked compartments (*hh*-Gal4-UAS-CD8-GFP) and morphological features revealed that small molecules are respecting classical and non-classical compartment boundaries
- Coherent anti-Stokes Raman Spectroscopy (CARS) micro spectroscopy indicated that lipids are differentially distributed within the wing imaginal disc with a relative higher content in the wing blade region compared to the body wall region
- Investigating the ToF-SIMS spectra allowed us to tentatively assign several diacylglyceroles to be more abundant in the wing blade region
- Generating lipid catalogues of wing imaginal discs enabled us to strengthen these assignments

Currently, we are in the process of further analyzing the wing imaginal disc with high end imaging mass spectrometry tools, in order to validate our previous findings.

The workflow in regard to the results presented herein opens up an entire new field in developmental biology. Future research will aim at clarifying the biological functions of molecules identified in this study.

Zusammenfassung

Die Entstehung eines Lebewesens aus einer einzelnen befruchteten Eizelle ist ein faszinierender Prozess. Um einen komplexen Organismus wie die Fruchtfliege *Drosophila melanogaster* zu generieren sind exakte Kontrollmechanismen nötig. Die Entwicklung des Körperplans mittels bedeutender Körperachsen hilft dabei die Entstehung zu steuern und zu kontrollieren. Hauptsächlich werden hierbei die antero-posteriore und dorsale-ventrale Achse unterschieden. Zusätzlich zum Gesamtorganismus sind die Hauptorgane ebenfalls anhand desselben Koordinatensystems organisiert. In *Drosophila melanogaster* wurde dieses Phänomen beispielsweise extensiv ausführlich am Vorläuferorgan des Flügels studiert. Die sogenannte Flügelimaginalscheibe ist in der antero-posterioren Achse durch die selektive Expression des Selektorgens *engrailed* strukturiert,, welches ausschliesslich in den posterioren Zellen exprimiert wird. In der dorso-ventralen Achse ist das Selektorgen *apterous* für die Organisation der dorsalen und ventralen Zellen zuständig. Die Expression von *apterous* ist hierbei ausschliesslich in dorsalen Zellen nachweisbar. Zusätzlich zu diesen klassischen Kompartimentsgrenzen ist die Flügelimaginalscheibe durch nicht-klassische Grenzen substrukturiert. Diese nicht-klassischen Grenzen sind meist durch morphologische Merkmale, wie beispielsweise Gewebefaltungen, sichtbar. Hierzu zählt vor allem die Unterteilung von zukünftigem Körpergewebe und dem zukünftigen Flügelblatt (B/W). Die oben aufgeführten Beispiele verdeutlichen, dass zur Entstehung von Kompartimentsgrenzen auf dem genetischen Level bereits viel erforscht wurde. Bisher wurde jedoch der gesamte Fundus von niedermolekularen Stoffen nicht analysiert. Daher untersuchten wir in dieser Arbeit mittels bildgebender Flugzeit sekundär Ionen Massenspektrometrie (ToF SIMS MSI), wie niedermolekulare Stoffe auf der Oberfläche der Flügelimaginalscheibe verteilt sind. Die Lokalisation entlang der klassischen und nicht-klassischen Grenzen wurde durch das Überlappen von fluoreszenz-markiertem posterioren Kompartiment und der Co-Registration anhand morphologischer Merkmale überprüft. Um die Identität der niedermolekularen Stoffe weiter einzugrenzen, wurde kohärente anti-Stokes Raman-Spektroskopie angewandt. Zur vorläufigen Identifikation der Stoffe, die lokal differentiell lokalisiert sind, wurde ein Katalog von Lipiden der Flügelimaginalscheibe generiert. Dieser Katalog wurde dann unter Einbezug der Literatur mit den ToF-SIMS MSI Daten verglichen. Zusätzlich wurden im Rahmen dieser Arbeit zwei Algorithmen entwickelt, die Aussagekraft verzerrter Bilder verbessern. Durch mehrere internationale Kooperationen konnten folgende Resultate erzielt werden:

- ToF-SIMS-MSI in positivem und negativem Ionisationsmodus zeigte, dass niedermolekulare Stoffe differentiell auf der Flügelimaginalscheibe verteilt sind
- Die Entwicklung zweier Bildkorrekturalgorithmen ermöglicht es partiell verzerrte Bilder zu rekonstruieren und interpretieren
- Hauptkomponenten (PCs) wurden mit markierten Kompartimenten (*hh*-Gal4-UAS-CD8-GFP) und morphologischen Merkmalen abgeglichen und konnten aufzeigen, dass niedermolekulare Stoffe die klassischen und nicht-klassischen Grenzen respektieren.
- CARS spektrale Mikroskopie weist darauf hin, dass Lipide in unterschiedlicher Anzahl zwischen dem zukünftigen Körpergewebe und dem Flügelblatt vorkommen. Bezeichnenderweise treten Lipide hierbei häufiger im Flügelblatt auf.
- Analyse der ToF-SIMS Spektren in Kombination mit Literaturrecherche führte zur Identifikation von unterschiedlichen Diacylglycerolen, die im Flügelblatt gehäuft auftreten
- Ein selbst erstellter Lipidkatalog der Flügelimaginalscheibe ermöglichte es uns die Identifikation der Diacylglycerole zu verbessern.

Mittels neuartigen bildgebenden massenspektrometrischen Methoden analysieren wir zur Zeit Imaginalscheiben., um unsere Ergebnisse zu validieren.

Der hier präsentierte Workflow in Bezug auf die von uns erzielten Resultate eröffnet ein vollkommen neues Forschungsfeld in der Entwicklungsbiologie. Die nachfolgende Herausforderung wird es nun sein, die biologischen Funktionen der von uns identifizierten Moleküle herauszufinden

Introduction

Developmental Biology

No process is more fascinating than the development of a single fertilized egg cell into a complete adult organism. Model organisms such as the fruit fly *Drosophila melanogaster* helped to identify many of the conserved key concepts of development (*1*). At first. cell division ensures there is the proper number of cells to take over all necessary tasks. Cell division is followed by pattern formation. At the very beginning this involves setting up the overall body plan. The main axis is established to determine the posterior (tail) and anterior (head) end of the body. Additionally, the dorsal (back) and ventral (front) axis is specified. These axes are maintained throughout the whole development. Moreover, most of the organs of the adult organism are also organized using the same coordinate system. Pattern formation goes along with cell differentiation to generate functionally specialized cells. Finally, after setting up the necessary structure the organism grows to its final size. Thus, development constantly continues through the entire life span of a multicellular organism.

Development of *Drosophila melanogaster*: A short recapitulation

Drosophila melanogaster has been studied for more than a century to gain knowledge about developmental processes. In this section, a brief overview of Drosophila development is presented, highlighting the features relevant to this thesis (*1*–*4*). The development of Drosophila can be subdivided in three very distinct periods. The first is embryonic development. After fertilization, the egg cell undergoes a series of rapid nuclear divisions without cell division. This leads to a syncytial structure with a common cytoplasm followed by cellularization and gastrulation. Embryonic development ends with the hatching of a larva. During three distinct larval stages (first instar, second instar and third instar), the organism develops further to set up all the structures necessary to produce the adult organism. During this period, all the major organs are set up as organ primordia called imaginal discs. Once larval development is completed, a pupa arises and metamorphosis takes place. Ten days after fertilization an adult fruit fly encloses from the pupa and the life cycle starts anew (Figure 1a).

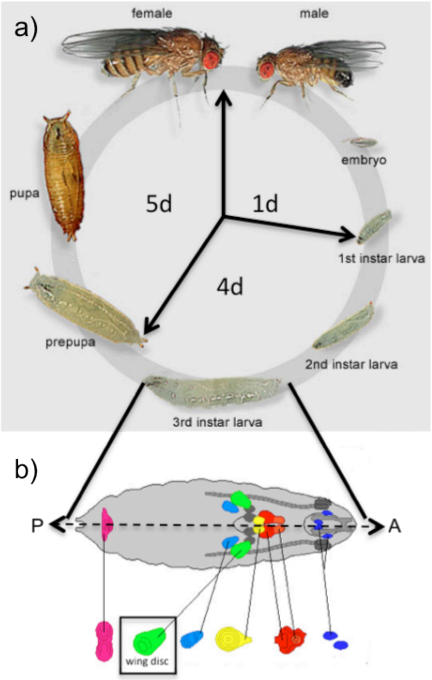
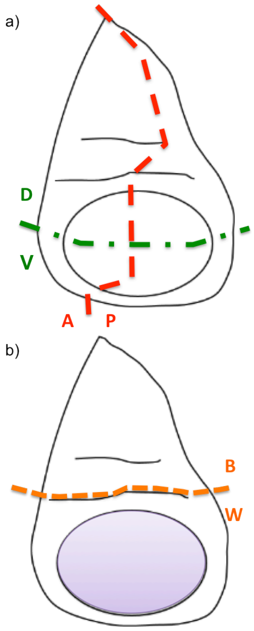


Figure 1: Development of *Drosophila melanogaster*: a) Life cycle of *Drosophila melanogaster.* 24 hours after egg laying embryonic development ends with a larva hatching. Within 4 days the larva undergoes two molts and finally reaches puparation. In the upcoming 5 days the pupa undergoes metamorphosis and a new adult fly hatches. **b)** During larval development the main structures of the adult fruit fly are set up as organ primordia called imaginal discs. For every major organ there is one imaginal disc set up. The imaginal discs are from P to A: Genital discs (purple), haltere discs (light blue), wing discs (green), leg disc (yellow), eye-anntenal disc (red), mouth parts (dark blue).

Development of the wing imaginal disc



Development of the wing(s) of *Drosophila melanogaster* has been used extensively as a model system to study organogenesis. During the first larval instar about 10-50 cells are predetermined to become the presumptive wing imaginal disc cells. Up to the third instar larva these cells undergo a tremendous amount of mitoses, leading to a final cell number of 30,000 to 100,000 cells (*5*) (Figure 1b). Analogous to the whole body the wing imaginal disc is organized along two main axes: the anterior-posterior (A/P) axis and the dorsal-ventral (D/V) axis (Figure 1 and Figure 2a). These axes are already set up in the embryo and are maintained through development and adulthood. Additionally, to these classical compartment boundaries the wing imaginal discs is also patterned by so-called non-classical boundaries (Figure 2b).

Figure 2: Classical and non-classical compartment boundaries in the wing imaginal disc: a) The wing imaginal disc development and growth is organized by two major classical compartment boundaries perpendicular to each other, namely the anterior-posterior (A/P; red line) and the dorsal-ventral (D/V; green line) boundary. **b)** Additionally, the disc is further patterned by other non-classical boundaries such as the body wall-wing blade (B/W; red line) or the pouch- (purple section) non-pouch tissue.

Formation of compartment boundaries in the wing imaginal disc

Classical compartment boundaries

Classical compartment boundaries are established in the early Drosophila embryo by the regulated activity of selector genes and are maintained through development and growth (*6*). In the wing imaginal disc the homeodomain-containing transcription factor Engrailed specifies posterior identity (*7*). Engrailed *(en)* leads to the transcription of *hedgehog* in the posterior compartment. Hedgehog diffuses in both directions. But only in a small stripe of anterior cells adjacent to the boundary the receptor *patched* is expressed. Binding of Hedgehog to Patched releases Smoothend to activate *decapentaplegic (dpp)* expression restricted to the stripe anterior of the A/P boundary. In the posterior cells, the *dpp* expression is repressed by *en*. This network ensures a proper positioning of the boundary during development and growth (*8–10*). Dpp then diffuses in both anterior and posterior direction to control development and growth in gradient manner. Moreover, cells expressing *engrailed* are terminally determined to be posterior and do not mix with cells from the anterior compartment. Even in the presence of cell division at the boundary, the posterior cells remain in their compartment and do not grow into the anterior compartment (*11*). Models ranging from different cell adhesion properties to the involvement of mechanical forces, such as tension, have been used to describe this behavior (*12*, *13*). Perpendicular to the A/P axis the (D/V) axis separates the dorsal from the ventral compartment. Here, the selector gene *apterous* is expressed only in dorsal cells and determines the identity of dorsal cells (*14*). Additionally, *wingless* production is controlled by bidirectional signaling of Serrate and Delta (*15*, *16*). *Wingless* is expressed in this way only in the cells of the dorso-ventral boundary. Secretion of wingless controls growth and patterning along the D/V axis of the wing imaginal disc (*17*, *18*). Thus, these two major compartment boundaries are responsible to control development and growth of the wing imaginal disc.

Non-classical compartment boundaries

In addition to the compartment boundaries established in a lineage restriction mode (classical boundaries), the wing imaginal disc is further subdivided into other compartments. Along these boundaries cells are restricted to the compartments because of differentiation rather than lineage restriction. Often, these boundaries are called "not-true compartment boundaries" or "non-classical boundaries" (*19*). The mechanism of differentiation often makes non-classical boundaries morphologically visible (*20*). One such boundary described in great detail is the boundary separating the future notum (body wall) from the wing blade region (further called B/W boundary) (*21*). It has been shown that this boundary is maintained by the expression of genes (*araucan, caupolican* and *mirror)* that encode proteins of the Iroquois Complex (Iro-C) and the mutual repression of *drop (msh* or *drp)*. The loss of members of the Iro-C transforms cells from notum to hinge if not later than the early third instar (*21*). Additionally, the apposition of cells expressing Iro-C members to cells not expressing the genes generates a cell sorting mechanism and leads to the fold appearing during the third larval instar (*22*, *23*). Generally spoken, the role of Iro-C in B/W patterning is similar to that of *engrailed* for the A/P and *apterous* for the D/V boundary formation. Importantly, after the major boundaries are established, it is crucial to maintain and control their positioning properly.

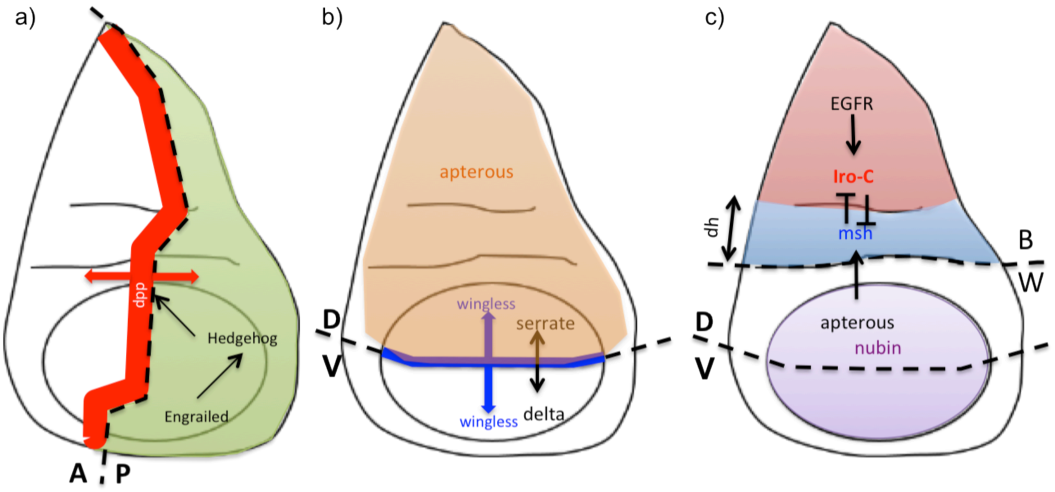


Figure 3: Genetics of classical and non-classical compartment boundaries of the wing imaginal disc: a) The A/P boundary (black dashed line) is established by the selector gene *engrailed* specifying posterior identity (green field). *Engrailed* leads to the expression of *hedgehog, which* diffuses in anterior direction. Through a set of genetic interaction this finally leads to the expression of *dpp* in a small stripe of anterior cells at the A/P boundary (red). *Dpp* then diffuses in both directions forming a gradient to control development and growth along the A/P axis (red arrows). **b)** Perpendicular to the A/P axis the D/V axis (black dashed line) is specified by the restricted expression of the selector gene *apterous* in the dorsal cells only (brown field). At the boundary, *serrate* and *delta* control the expression of *wingless* in a small stripe (blue). In a bidirectional gradient, *wingless* controls development and growth of cells in the D/V dimension (blue arrows). Both the A/P (a) and the D/V (b) boundaries are lineage restriction boundaries. **c)** Beside the classical compartment boundaries, the wing imaginal disc contains other non-classical compartments. The body wall is separated from the wing blade region (more dorsal dashed black line) by the restricted action of members of the Iroquois Complex (Iro-C) (red field), which is under the control of Epidermal growth factor receptor (EGFR). Within the body wall part the disc is further subdivided into the dorsal hinge (dh, blue area) and the notum (red area) by the activity of *msh* which is mutually exclusive to Iro-C and controlled by *apterous*. During third larval instar, this leads to two folds appearing in the wing imaginal disc. A further compartment is specified by the expression of *nubin* restricted to the wing pouch region (purple area), which gives rise to the wing blade during metamorphosis.

Maintenance of compartment boundaries

As described above, many of the genes involved in the process of pattern formation and establishment of compartments have been identified and studied. Yet, much remains to be elucidated. How boundaries are maintained throughout the different stages of development is not completely understood. Clearly, a sharp boundary needs to be maintained during development in order to ensure proper organogenesis. In the very early stages of boundary biology the differential cell adhesion theory was established (*24*). The assumption for this theory is based on the idea that neighboring cells from different compartments show a different adhesive tendency for each other. This is responsible for cell sorting along the boundaries on a cadherin basis (*25*). Additionally, a recent study claims that accumulation of extracellular matrix (ECM) is responsible for boundary formation between rhombomers in chicken embryo (*26*). A similar mechanism might also play a partial role in boundary maintenance in Drosophila*.* As cell division will challenge the sharpness of the boundary, it is likely that reduction of cell division may help to define and maintain the boundaries. Indeed, it has been shown that the rate of proliferation is reduced near the D/V boundary. If this phenomenon is truly responsible for the boundary maintenance remains to be clarified (*27*, *28*). Another mechanism involving mechanical forces is also assumed to play a role in maintaining the compartment boundaries. It has been described that along the major axes (A/P and D/V) F-actin and non-muscle myosin II are enriched. The authors of this work speculate that this enrichment leads to an increase of mechanical tension along the boundary in order to sharpen it (*13*, *29*).

By itself, none of the mechanisms could describe compartment boundary maintenance completely. Likely, it involves the interplay of many of the known aspects, as well as some yet undiscovered factors. Strikingly, the plethora of literature about compartment boundary formation and maintenance has only dealt with genes and their direct products (e.g. RNA, proteins). To our knowledge, the involvement of small molecules, such as lipids, in the previously described processes has not yet been studied. Since lipids are the key molecules in membranes of biological systems, it is likely that the lipid composition has a strong influence on the properties of cells. One reason that lipids have not been studied in this context so far is because of the lack of proper analysis tools until recently. Technologies providing chemical specificity combined with imaging capabilities are needed to investigate the role of small molecules in the process of pattern formation and maintenance. The recent development of mass spectrometric imaging (MSI) is able to close this gap and may lead to new discoveries in the biology of compartment boundaries.

Imaging mass spectrometry

General principle of mass spectrometry

Every mass spectrometer consists of the same three units: an ionizer, an analyzer and a detector to measure the ions (*30*). However, a large range of biomolecules was not accessible due to lack of "soft ionization" methods in the beginning of utilizing mass spectrometry. With the application of electrospray ionization (ESI) or desorption/Electrospray Ionization (DESI) in MSI experiments as well as matrix-assisted laser desorption/ionization (MALDI) this gap was closed (*31*, *32*). Together with secondary ion mass spectrometry (SIMS) the whole range from atoms over small molecules (e.g. lipids) up to several hundreds of kDa heavy protein complexes can now be routinely analyzed and quantified. But not only the ionization process was and still is under massive development. Also, on the analyzer part much has been improved since the first ToF analyzers. The introduction of ion traps, quadrupoles, Fourier Transform (FT) and Orbitrap analyzers has increased the spectral resolution, the mass accuracy, the sensitivity and the amount of ions that can be detected in one run (*33*). Last but not least, also the detector field is still under development (*34*). Combination of multiple individual parts allows the generation of hybrid mass spectrometers to fulfill entirely different tasks (*33*). This catalyzes the ongoing improvement of instrumentation to reveal multidisciplinary (chemistry, physics, biology) phenomena. Moreover, the need for technologies able to combine spatial information with the chemical assignment provided by general mass spectrometry led to the development of the field of imaging mass spectrometry.

Figure 4: General principle of mass spectrometry: Every mass spectrometer consists of three units: First, an ionizer (orange box) which desorbs and ionizes the analyte; following ionization, the ions are analyzed in an analyzer (blue box) based on mass to charge (m/z) ratio; last, the ions are detected and recorded on the far end of the instrument on the detector (green box).

General principles of imaging mass spectrometry

By performing mass spectrometry in imaging mode, many organic and inorganic molecules can be detected and localized simultaneously, even from the most complex surfaces, such as biological tissues. A complete MSI run involves sample preparation, desorption/ionization, mass analysis, followed by detection of the ions, and the registration of the images post data acquisition (*35*). Therefore, MSI can also be described as position specific mass spectrometry. Taking advantage of the developments in MS, the field of MSI is evolving rapidly. Instruments with high spatial resolution (lateral resolution) and/or high mass spectral resolution have been developed to make MSI the only known technology to generate images of biological relevant molecules directly from tissues in a label-free manner (*36*–*39*). Without much sample preparation the sample to be analyzed is introduced in a mass spectrometer. A mass spectrum is recorded at a first well-defined position with a defined spatial resolution. The mass spectrum is stored together with its coordinates. Once the first cycle is complete, the ionization source is moved to the next tile and a new spectrum is recorded and stored. This is repeated in predefined manner over the entire region of interest (ROI). Post data acquisition, the data analysis and interpretation tools allow the generation of molecular maps of the surface of interest. Although MSI profits from developments in the entire mass spectrometry community, only secondary ion mass spectrometry (SIMS) is further described. A thorough explanation of MALDI and DESI with its applications can be found in literature (*35*, *40*, *41*).

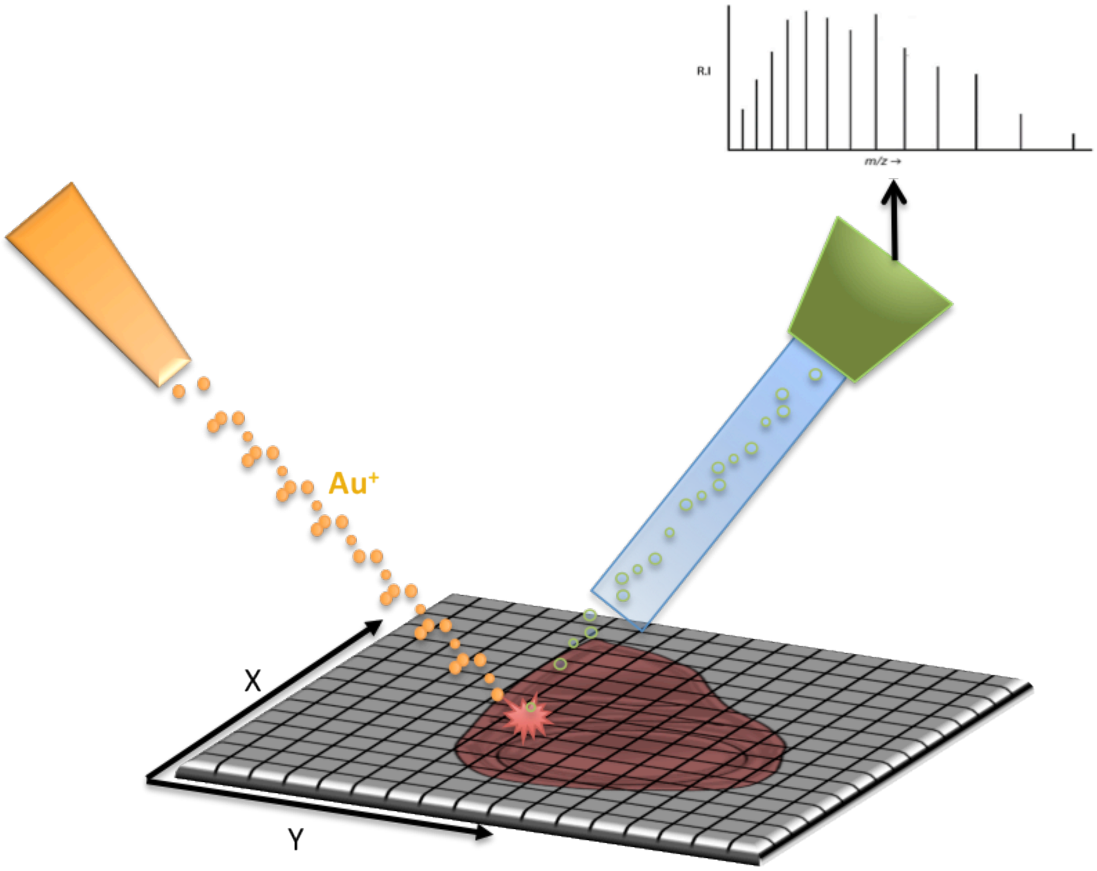


Figure 5: Principle of imaging mass spectrometry: Imaging mass spectrometry can be described as position specific mass spectrometry. The sample of interest (SOI) is placed onto a (conductive) glass support (gray field). The ionization source (orange) is focused onto a defined spot size for analysis. In this study, we used a liquid metal ion gun (LMIG) with Au^+ as primary ions (orange dots) to generate secondary ions (green circles) from the wing imaginal disc surface. Spot size of the primary ions determines the spatial resolution. After analysis (blue) and detection (green) of the ions, the spectrum is stored with its x,y-coordinates. Subsequently, the beam is moved to the next position and the procedure is repeated until the entire region of interest (ROI) is measured.

Secondary ion mass spectrometry imaging (SIMS MSI)

SIMS imaging mass spectrometry was among the very first technologies used to inorganic and organic surfaces. In the very beginning of its development it was mainly used to study surfaces of semiconductors and similar materials (*42*). In SIMS a primary ion beam with high keV energies is used to generate secondary ions from the surface of interest. Using appropriate primary ion beams and settings, SIMS is able to achieve spatial resolution as high as 50nm, with secondary ions being emitted only from the first 5-10nm of surface. Unfortunately, the ionization efficiency of SIMS is rather low with only 0.1% of the emitted secondary particles being ionized (*43*). The vast majorities of the molecules are released as neutrals and cannot be detected. Moreover, compared to MALDI and DESI, the energy of the ionization beam is much higher. In order to prevent extensive fragmentation, the primary ion dose has to be kept very low. At a dose below 10^{13} primary ions*cm^{-2} it is statistically unlikely (<1%) that the same area of the surface is sampled twice to minimize fragmentation (*44*, *45*). Therefore, the information gathered this way reflects the chemical composition of the surface (static SIMS). Above this value (the static limit of SIMS) extensive fragmentation takes place. To deal with the drawbacks described above, researchers are working on different primary ion sources to facilitate imaging of biological surfaces and improve the ionization efficiencies. One first breakthrough was the development of liquid metal ion guns (LMIG) with monoatomic primary ions such as Au^+ (*46*). Further developments lead to the use of polyatomic primary ion sources for SIMS. Especially C_{60}^+ displayed very favorable features. Foremost, polyatomic primary ions desorb the surface with significantly higher ion yields and less fragmentation of the secondary ions emitted. Therefore, intact biomolecules can be analyzed even better (*36*). Additionally, surface modifications have been developed to further enhance the ionization of intact biomolecules. Routinely, a thin layer of metal, gold or silver, is applied. This technique is referred to as metal-assisted SIMS (MetA-SIMS). Further, the use of matrices as in MALDI was also examined for SIMS (matrix-enhanced (ME) SIMS). Applying these technologies, the sensitivity for larger ions is tremendously increased. This allows the detection of intact lipids, fatty acids, and other molecules up to a size of 5kDa (*47*, *48*). Also, the sample metallization makes biological surfaces conductive. Normally, they are considered insulators and are therefore not able to be analyzed without charge compensation (*46*). Potentially, the coating of surfaces with metals leads to secondary ions that are often adducts of the coating metal and the endogenous molecules. This can make the interpretation of the spectra slightly more difficult as the observed peaks are no longer from the molecular ion but different kind of adducts.

The ToF analyzer is the most widely used mass analyzer in SIMS MSI. In a ToF analyzer the ions are separated in relation to their flight time through a defined length (*49*). To ensure proper mass analysis it is crucial that all ions are accelerated with the same kinetic energy. In this way, the ions arrive at the detector at different times proportional to the square root of the m/z value depending on their initial energy (U) and the flight path length (d).

$$t = \frac{d}{\sqrt{2U}} * \sqrt{\frac{m}{q}}$$

With $\frac{d}{\sqrt{2U}}$ being constant (k) for a single analyzed pulse of ions leading to

$$t = k * \sqrt{\frac{m}{q}}$$

With m being the mass and q=1 the charge for SIMS. ToF analyzers offer a very high ion transmission rate with good sensitivity and good dynamic mass range. After the ions pass the mass analyzer they hit a detector at the other end of the instrument. Most commonly a detector consists of a micro-channel plate where the ions are converted to electrons followed by a scintillator to convert the electrons into photons. Outside the vacuum a photomultiplier is placed to record the events hitting the detector. All information is stored in a matrix fashion with m/z and intensity (I) at every position (x,y).

Mass analyzers with higher accuracy, better sensitivity, and increased spectral resolution are used more and more often in SIMS. Recently, Smith and colleagues (*36*) developed a unique system combining a C_{60} primary ion gun with a Fourier Transform Ion Cyclotron Resonance (FT-ICR or FT) mass analyzer. In addition to the features described above, a FT-ICR analyzer allows tandem MS(MS/MS) experiments. MS/MS is especially important to unambiguously identify the imaged molecules (*50*). Furthermore, a FT analyzer can be combined with a wide variety of primary ion sources like MALDI and ESI.

Data analysis strategies

Mass spectrometric imaging data sets contain hundreds of thousands m/z values with the corresponding intensities at every x,y-position recorded. The data load and the data complexity make it almost impossible to manually interpret them. Due to its structure, however, the data can be interpreted as a matrix and analyzed by different statistical tools (*51*). Commonly, the data is analyzed using a univariate or a multivariate analysis method. In a univariate analysis strategy the software provides a distribution over the entire sample of individual peaks. Most software tools available allow the rapid browsing through entire peak lists to screen for interesting distributions of known and unknown molecules (*52*). Still, data analysis using these strategies is very time consuming. Therefore, multivariate analysis strategies have been developed to further improve data analysis. If every measured pixel is considered an individual sample, a score plot in the spatial dimension can be generated as an output of the multivariate methods. A more detailed overview of these methods can be found in (*51*). For the scope of this work, only principal component analysis was used and is therefore described further.

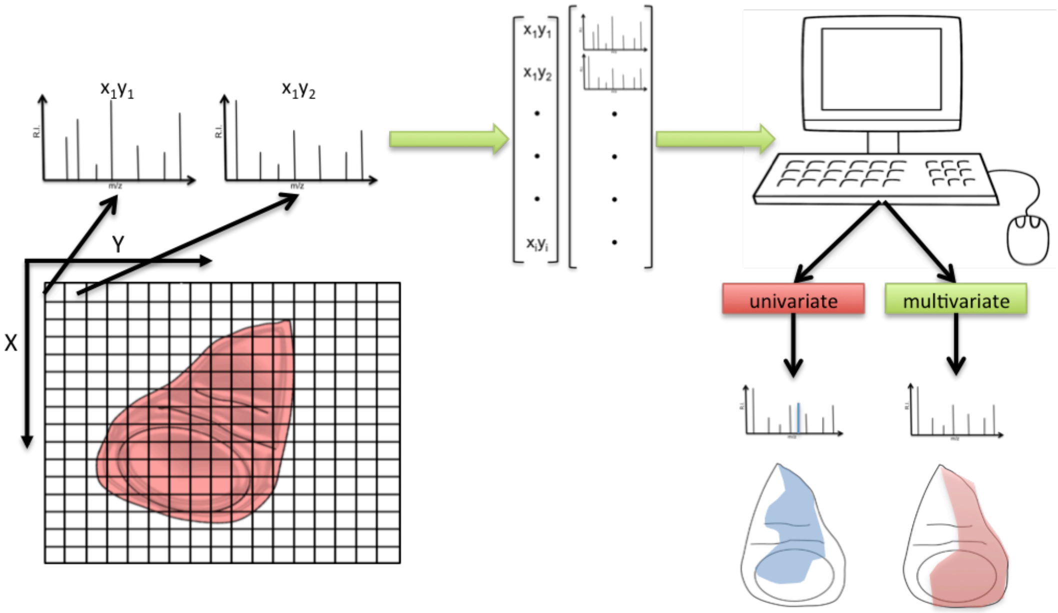


Figure 6: Data structure and data analysis overview: In an imaging mass spectrometry run a mass spectrum is recorded at every x,y-position. The intensity of the m/z values is stored together with the location (x.y). The data can be summarized as a matrix containing the x,y-positions and the intensities of every m/z value at a specific position. Data processing and analysis allows the visualization of individual peaks (m/z values) over the region of interest (ROI) in an univariate mode (red text box). Additionally, multivariate strategies (green text box), such as principal component analysis (PCA), can be applied to analyze the complex data sets. These strategies allow the researcher to draw conclusions about the distribution of m/z values over the tissue of interest.

Principal component analysis (PCA)

Initially used in the field of secondary ion mass spectrometry, today PCA is the most widely used multivariate data analysis technique in the field of MSI. Simplified, PCA can be described as a tool for measuring relationships between variables and samples with peaks being the variables and pixels being individual samples (*53*, *54*). It tries to minimize the number of variables used to describe a multidimensional data set. Importantly, PCA aims to keep the loss of information at a minimum. In this way, the variance is used to distinguish different groups and transform them into principal components (PC). The PCs are ordered in decreasing variance from PC1 to PCn. Often, the first PC contains most of the variance of the entire dataset. To visualize the output of PCA score plots, loadings spectra, and the variance are typically presented. PCA is a very efficient and fast algorithm to extract most key features of a data set. Often, the difference between the sample (tissue) and the sample holder is described in the first PC. In subsequent PCs remaining differences between several regions of the investigated tissue are displayed (*54*–*56*). Apart from PCA, many other multivariate techniques, such as clustering methods (e.g. k-means clustering) or factorization methods (e.g. non-negative matrix-factorization (NNMF)), have been developed; they are reviewed extensively in (*51*).

Limitations of MSI

Sample preparation is the most crucial step for every MSI experiment. Importantly, sample preparation needs to take place under defined conditions to prevent degradation of molecules. It is reported that degradation of some molecules can already appear within few minutes (*57*). Fortunately, for SIMS the sample pretreatment is not very time consuming. After mounting the sample onto conductive slides they can be either coated with a thin layer of metals (MetA-SIMS), covered with a matrix similar to MALDI MSI, or directly introduced into the instrument. The complexity of biological tissues is another challenge. Molecules distributed over the tissue can enhance or prevent ionization of a certain molecular class. Additionally, ion suppression due to the presence of salts is a major concern. This can lead to ion images which do not reflect the true distribution of a certain species (*58*, *59*). Also, salts present on the surface as well as the metallization of the sample can lead to adduct formation of the secondary ions. This can hamper spectral investigation and assignment of peaks to molecules.

Advantages of MSI

Although many challenges have to be overcome during an MSI run, the technology still offers a multitude of features that make it unique in the bio-analytical field. First and foremost, MSI couples the ability to chemically identify molecules via MS with microscopic capabilities. Further, it can detect a wide range of molecular classes (e.g. lipids, peptides, proteins) at very high spatial resolution. Development of very high mass accuracy instruments improves the assignment of peaks observed as the molecular formula of the peak can be determined based on the accurate mass assignment of the peak. Additionally, the improved sensitivity of these instruments enhances the number of detected ions per measurement. Further, MSI is completely label-free with easy and fast sample preparation (*60*). Altogether, this makes MSI a well-suited tool to screen the surface of the *Drosophila melanogaster* wing imaginal discs for small molecules involved in pattern formation and maintenance of compartment boundaries.

Project rational

The project presented herein targets the examination of the spatial distribution of small molecules in the wing imaginal disc. This examination was carried out with secondary ion mass spectrometry imaging (SIMS-MSI) and other surface analysis techniques (microscopy, CARS spectroscopy).

To our knowledge, for the first time the potential involvement of small molecules (e.g. lipids) is analyzed. SIMS-MSI with its chemical specificity combined with microscopic imaging capabilities enables direct analysis of molecules of interest on the tissue surface. In addition, SIMS-MSI is a label-free technology, which makes sample preparation straightforward. Alternative strategies are used to generate a map of small molecules on the wing imaginal disc with unambiguous identification of these small molecules to verify the results from SIMS-MSI differently.

The main goals/objectives are:

- Set up SIMS-MSI protocols including sample preparation to measure the distribution of small molecules
    - Several factors such as tissue topography, tissue integrity, and prevention of degradation need to be taken into account in order to ensure correct measurements
- Measure third instar wing imaginal disc in positive and negative ion mode on a SIMS-TRIFT-ToF equipped with a liquid metal ion gun (LMIG) with Au^+ as primary ions to test suitability
- Perform univariate and multivariate (PCA) statistical analysis to identify patterns of interest or new patterns in an unbiased way
- Correlate identified patterns with known markers using the UAS-Gal4 system available in Drosophila, as well as morphological features
- As ToF-SIMS MSI lacks unambiguous identification, it is crucial to apply further hyper-spectral technologies to define the possible nature of small molecules
    - Herein, we used coherent anti-stokes Raman (CARS) micro spectrometry to obtain information about the molecular classes
- Use literature and generated lipid catalogues with liquid chromatography MS/MS (LC-MS/MS) to assign lipids to different compartments
- Apply high mass accuracy, high spatial resolution SIMS MSI to identify the lipids of interest directly on tissue with C_{60} FT-ICR and/or MALDI FT-ICR MSI

The focus of this thesis was to evaluate if small molecules on the surface of the wing imaginal disc exist that have a spatial distribution similar to the genetically established compartments. Functional analysis of the identified molecules is beyond the scope of this work, but is a main topic for future studies.

Importantly, the study presented here is a multi-national multi-disciplinary effort. Many people in many different laboratories have helped to generate the results presented. Crucial for ToF-SIMS MSI was the collaboration with the Biomolecular Imaging Mass Spectrometry (BIMS) laboratory of Prof. Ron M.A. Heeren at the FOM Institute AMOLF, Amsterdam, NL. During a longer stay in the BIMS group, Florian Marty set up several other collaborations with the help of Dr. Erich Brunner, Prof. Konrad Basler and Prof. Ron M.A. Heeren. These collaborations are

- CARS spectro microscopy with Gianluca Rago of the Mischa Bonn group at the FOM Institute AMOLF, Amsterdam, NL
- C_{60} SIMS FT ICR MSI with Don Smith of the BIMS group of the FOM Institute AMOLF, Amsterdam, NL, and Franklin E. Leach and Christopher S. Anderton from the laboratory of Dr. Ljiljana Pasa-Tolic at the EMSL/PNNL, Richland, WA, USA
- LC-MS/MS analysis of lipid extracts from third instar wing imaginal discs in collaboration with Dr. Xiaoli Gao from the Department of Biochemistry University of San Antonio, TE, USA
- Mathematical correction of distorted imaging mass spectrometry data sets in collaboration with Florian Gerber from the group of Prof. Reinhard Furrer, Institute for Mathematics, University of Zurich and Gert Eijkel from the BIMS group, FOM Institute AMOLF, Amsterdam, NL

These collaborations led to a first author publication for Florian Marty during his PhD studies accepted in Analytical Chemistry and a second first author manuscript to be submitted. Additionally, a co-author publication was obtained with Bjoern Handke from Prof. Christian Lehners group, Institute of Molecular Life Sciences, University of Zurich.

The manuscripts of the two first author publications are presented in the results section. The co-author publication is published in PLoS One under the running title *"The hemolymph proteome of fed and starved Drosophila larvae"*.

Results

Imaging mass spectrometry detects small molecules respecting the compartment boundaries on the wing imaginal disc

Small molecule distributions in the *Drosophila melanogaster* wing imaginal disc revealed by ToF SIMS imaging mass spectrometry and CARS spectromicroscopy

Authors: Florian Marty[1$], Gianluca Rago[2,3$], Donald F. Smith[2], Xiaoli Gao[4], Gert B. Eijkel[2], Luke MacAleese[2†], Mischa Bonn[2,3], Erich Brunner[1], Konrad Basler[1*], Ron M.A.Heeren[2*]

Affiliations:

[1]Institute of Molecular Life Sciences, University of Zurich, Winterthurerstrasse 190, CH-8057 Zurich, Switzerland

[2]FOM-AMOLF, Science Park 104, 1098 XG Amsterdam, The Netherlands

[3]Max Planck Institute for Polymer Research, Ackermannweg 10, 55128 Mainz, Germany

[4]Institutional Mass Spectrometry Laboratory, The University of Texas Health Science Center at San Antonio, 8403 Floyd Curl Drive, MC-7760 San Antonio, TX, USA

*Correspondence to: r.heeren@amolf.nl, konrad.basler@imls.uzh.ch

†Current addresses: GR, LM, AK, MB

[$]Authors contributed equally to the work

One Sentence Summary: Hyper spectral imaging technologies reveal the inter-organ distribution of small molecules

Introduction

Living matter consists of a complex arrangement of molecules. It is one of the fundamental goals of biology to analyze, describe and ultimately understand the spatial distribution of biomolecules in relation to biological function. In contrast to the wealth of knowledge that has been obtained about the cellular arrangement of genes, the corresponding RNAs and proteoforms, the organization and distribution of small molecules such as carbohydrates and lipids is less well understood. This lack in our knowledge can be traced to the challenges associated with the localization of small molecules with high spatial resolution as, for instance, fluorescent labels are prohibitively large. Recent studies in the model system Drosophila melanogaster have described and quantified lipids over the course of its life cycle. Although these studies have provided important insights, demonstrating e.g. that different tissues exhibit distinct phospholipid compositions (*1*, *2*), the intra-organ distribution of small molecules has not been reported to date. Such insights into the detailed spatial distribution of proteins, lipids and nucleic acids are crucial to be able to relate specific molecular moieties to specific has often been crucial to place them into the appropriate functional context studying developmental processes or disease models, for instance in simple organ systems such as the imaginal discs of Drosophila melanogaster. The imaginal disc is [brief description] and constitutes a well-studied model system for [brief background].

In this report we reveal the distribution of small molecules in the Drosophila wing imaginal disc, using a combination of ToF SIMS imaging mass spectrometry and coherent anti-Stokes Raman spectroscopy (CARS) micro spectroscopy. These two complementary, label-free techniques make use of inherent properties of small molecules, namely their mass and their vibrational properties. We demonstrate that small molecules show well-defined distributions within this tissue. Strikingly, these distributions strongly resemble genetically established patterns known to be essential for wing development and growth. The spectral components from CARS reveal a lipid origin for some of these molecules, which were confirmed by the assignment of the ToF SIMS mass spectra to distinct diacyl- and triacylglycerol species. We validated the assignments using a reference catalogue of membrane lipids generated from wing imaginal discs. The discovery of regional differences raises the notion that small molecules play previously unrecognized roles in organ development.

Results:

To reveal the distribution of small molecules, third instar wing imaginal discs were measured with ToF SIMS imaging mass spectrometry. ToF SIMS was first applied in negative ion mode (*3*). However, we were unable to detect specific distribution patterns of single molecules by univariate data analysis strategies. It has been shown previously that tissue subtypes can be distinguished based on the profiles of multiple molecules rather than on single molecules (*4*). Therefore, we applied a multivariate statistical analysis method called principal component analysis (PCA) to analyze the ToF-SIMS data sets. We investigated the principle components (PC) for distribution patterns. Surprisingly, we identified PCs reflecting distributions reminiscent of known developmental patterns (*5–7*). The PCs identified discriminate anterior from posterior (A/P) (**Fig. 1a**) and dorsal from ventral tissue (D/V) (**Fig. 1b**), respectively. Additionally, patterns reminiscent of the body wall-wing blade differentiation (B/W) (**Fig. 1c**) and the pouch/non-pouch (**Fig. 1d**) patterns were observed. Performing the same workflow in positive ion mode we identified only the B/W distribution (**Fig. 1e**). We did not observe any of the other patterns seen in negative ion mode. Remarkably, we did not observe any unknown patterns in any of the ionization modes. All the principle components (PC) reflected either known patterns or homogeneous distributions across the entire tissue.

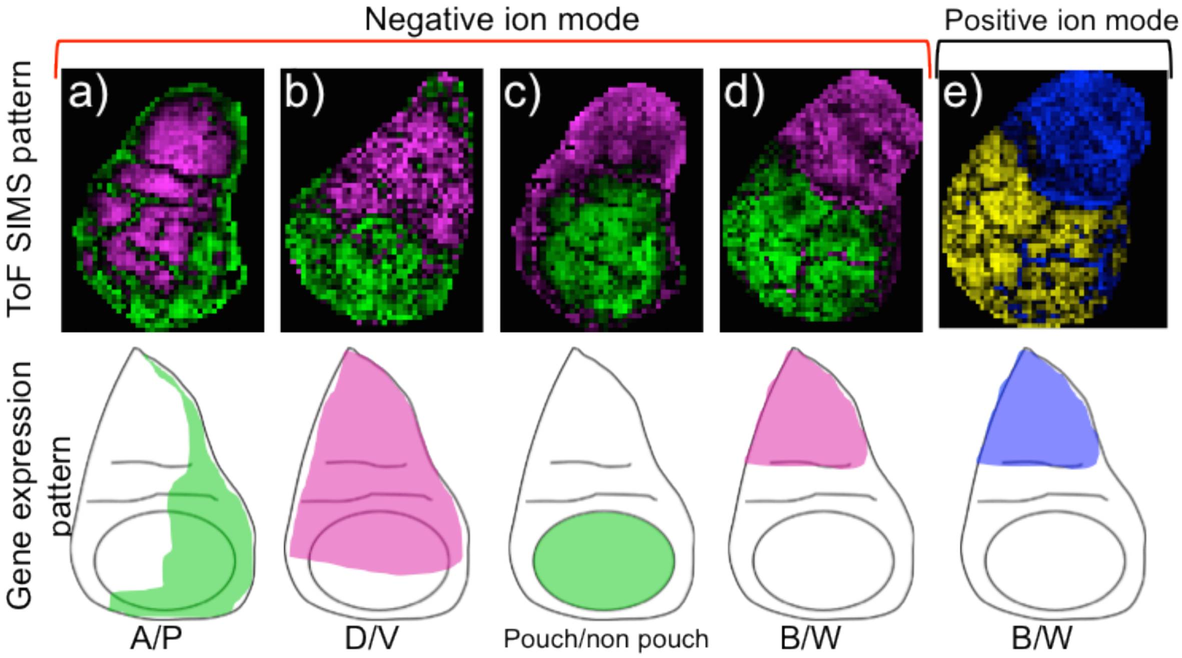


Fig. 1. Small molecule distribution revealed by ToF SIMS imaging mass spectrometry on the wing imaginal disc and corresponding genetically established patterns: ToF SIMS IMS and PCA revealed patterns reminiscent of the compartmental A/P (**a**) and D/V (**b**) subdivisions. The non-classical compartments pouch/non-pouch (**c**) and body-wall/wing-blade (B/W) d) were also detected. All these compartments were only observed in negative ion mode. In positive ion mode only the B/W subdivision could be observed (**e**).

To further validate that the PCA outputs reflect the known compartment boundaries, we imaged a genetically modified disc expressing green fluorescent protein (GFP specifically in the posterior compartment (**Fig. 2a**). Overlay of the ToF SIMS PCA (**Fig. 2b**) output with the fluorescent image revealed that the subdivisions observed by ToF SIMS are similar to the compartmentalization established by known genetic circuits (**Fig. 2c**). Additionally, we compared the B/W pattern of the ToF SIMS PCA image with the total ion image of the measurement. The B/W pattern is morphologically visible as a fold appearing in late third instar stage (**Fig. 2d**); at this fold the compartment boundary is established (*8*). As our measurements took place under high vacuum the fold expands to a crack and can be identified in the total ion image (**Fig. 2e**). This crack is exactly where the PCA differentiates between the two subdivisions (**Fig. 2f**). These findings indicate that small molecules are differentially distributed on the wing imaginal disc and may contribute to developmental processes such as pattern formation and compartment boundary maintenance.

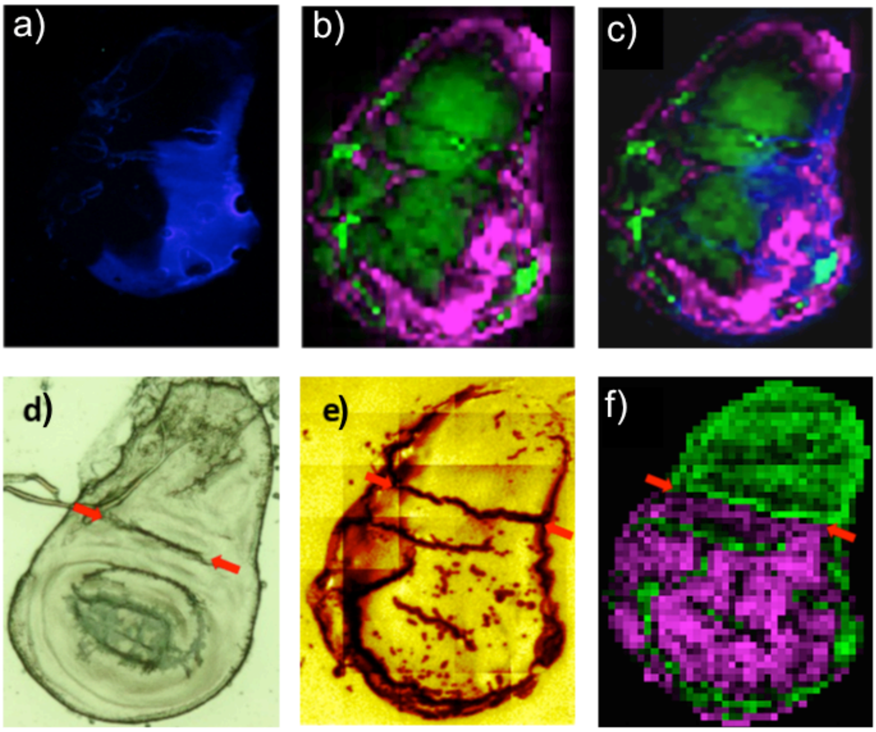


Fig. 2. Small molecule patterns identified by ToF SIMS imaging mass spectrometry mimic genetically established patterns: In order to examine whether the anterior-posterior subdivision coincides with the classical compartment boundary established by differential expression of *engrailed* (*en*) and *hedgehog* (*hh*), we generated fluorescent images of imaginal discs expressing GFP under the control of *hh* (*9*) (**a**). The PC representing this subdivision (**b**) was then overlaid with the fluorescent image (**c**). We used the fold appearing in late third instar (**d, e, f**, red arrows) as a marker for the non-classical B/W pattern (**f**). This fold expands to a crack in the high vacuum of the instrument without separating the tissue (**E**, red arrows). The PC distinguishing body wall from wing-blade exactly changes at this fold (**F**, red arrows). Thus, both boundaries revealed by PCs coincide with genetically established boundaries.

Due to the complexity of the ToF SIMS spectral data it was not possible to focus the subsequent analysis on a particular subclass of small molecules such as lipids, carbohydrates or peptides **(Fig. S1)**. All of these classes are composed of a very large number of single molecules. Therefore, it is crucial to identify the main molecular classes contributing to the patterns observed in the PCA of the ToF SIMS. The implementation of CARS and the chemical specificity it provides allowed us to assign the molecular class to certain patterns. CARS is a non-invasive microscopy approach where image contrast is provided by molecular vibrations that are inherently different for different molecules. We acquired the vibrational signatures of the molecules (in the frequency range corresponding to Raman shifts from -3200cm^{-1} to -900cm^{-1}) from late third instar wing imaginal discs. In the fingerprinting region (-1800cm^{-1} to -900cm^{-1}), the chemical moieties giving rise to specific peaks is largely known and can be found in (*10*). Because of its inherent confocal implementation, CARS micro-spectroscopy allows the generation of 3D topological maps based, for instance, on the C-H stretching region -3020cm^{-1} to -2850cm^{-1} **(Fig. S2)**. This capability is important as the topography of the sample could lead to artificial spectral features in ToF SIMS imaging mass spectrometry (*11*). Further, we ensured that the sample integrity is maintained **(Fig. S2)**. Multivariate statistical analysis (PCA) was applied to screen for patterns in the CARS dataset. In the second PC we observed a compartmentalization similar to the B/W observed in ToF SIMS **(Fig. 3a)**. Apart from this, we did not detect any other known or unknown patterns similar to the results observed in ToF SIMS imaging mass spectrometry. In a next step we performed canonical cross correlation analysis (CCA) to test how the different patterns observed with ToF SIMS imaging mass spectrometry and CARS overlap. Clearly, the B/W-like pattern observed applying both methods shows a significant correlation score **(Fig. 3c, d and Table 1)**. Thus, the chemical signature in the CARS PC may reveal the potential chemical components giving rise to the pattern observed in ToF SIMS. Therefore, we have investigated the spectral region of the CARS measurement from the correlating value 2 (CV2) **(Fig 3b).**

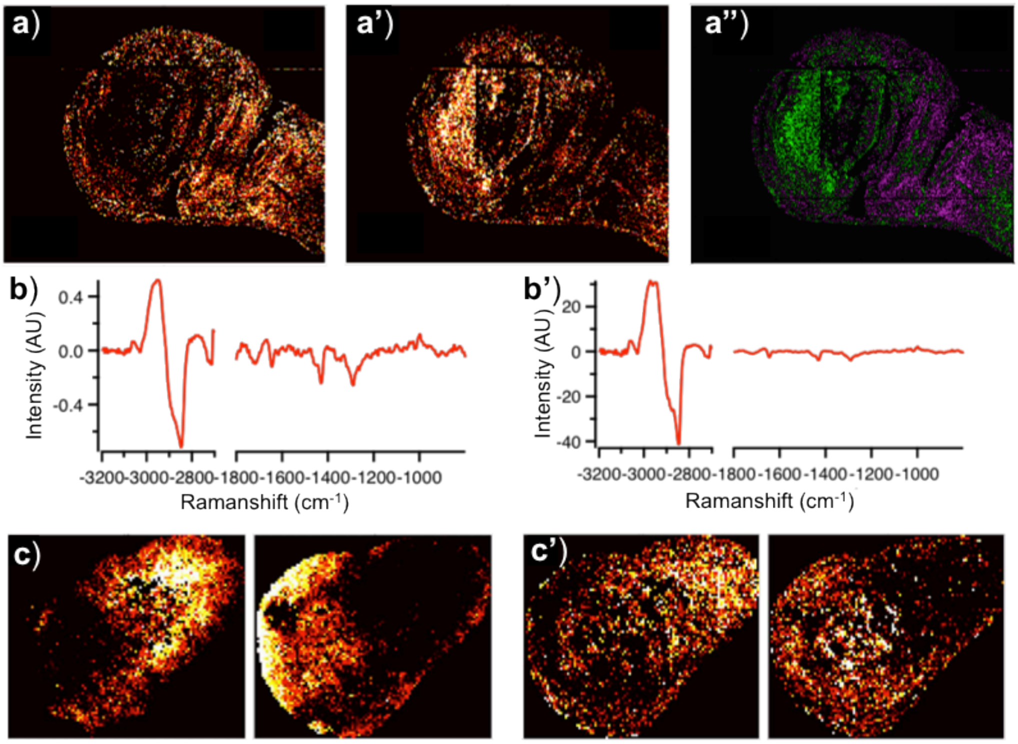


Fig. 3. Coherent anti-Stokes Raman spectroscopy (CARS) reveals a lipid origin for the B/W pattern: CARS was applied to test if patterns observed by ToF SIMS can be reproduced and to gain further knowledge about the molecules involved. The B/W pattern could be demonstrated with PCA (**a, a', a''**). Investigation of the normalized spectra revealed a higher content of lipids in the wing blade region (**b, b'**). Images of wing imaginal discs were then analyzed with both PCA and CCA methods and canonical cross correlation was applied to test the correlation of the observed patterns from ToF SIMS (**c**) and CARS (**c' and Table 1**).

		CARS	SIMS
CV	Correlation	Variance described	
1	0.394264	4.984	0.69102
2	0.288557	3.600	0.54064
3	0.057636	3.209	0.50543
4	-0.048119	9.538	0.97901
5	-0.020349	7.359	0.34812
6	-0.021041	1.867	0.72842
7	-0.005737	2.181	0.39812

Table 1. Canonical cross correlation scores for patterns observed in ToF SIMS and CARS. Correlation was examined with two independent methods, both indicating that the first two CV do significantly correlate.

The dispersive C-H feature and the phenylalanine peak at -1000cm^{-1} point to a protein dominance for the body wall part, whereas the spectral components indicate a dominance of lipids in the wing blade section. To try to elucidate the nature of the lipids we investigated in more detail the PCA components associated with the B/W pattern. In independent ToF SIMS measurements we could detect the same pattern repetitively (**Fig. 4, Fig. S3**).

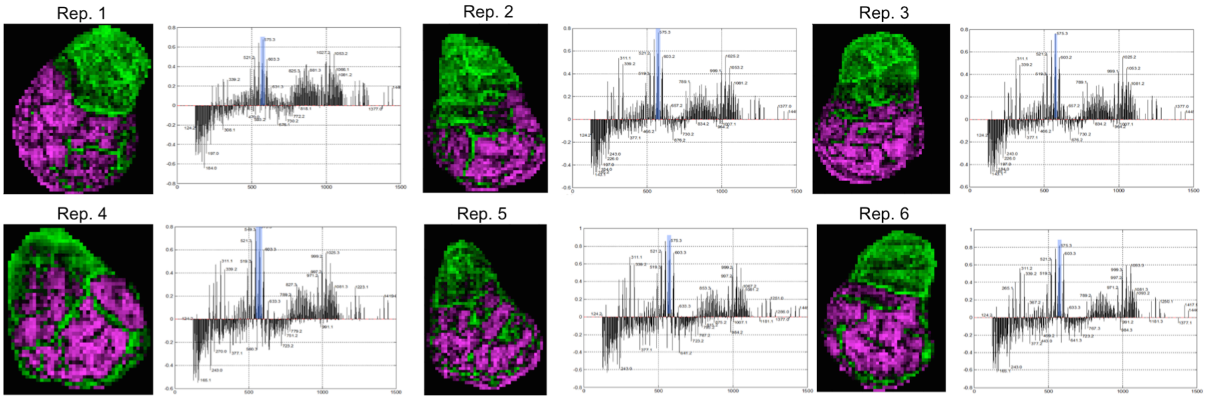


Fig. 4. Repetitive B/W pattern observed in positive ion mode: ToF SIMS analysis of a series of late third instar wing imaginal discs revealed the B/W pattern multiple times identified by PCA. Besides the score plot images the spectra resemble the B/W pattern as well.

We observed a highly reproducible pattern of peaks. Based on previously assigned m/z values we identified certain peaks using the spectral information of the positive ion mode (**Fig. 5**) (*12*). This approach allowed us to tentatively assign 22 lipids to the classes of monoacyl- (MG), diacyl- (DG) and triacylglycerol (TG). Further, in order to validate the identification of the ToF SIMS peaks we performed orthogonal validation using lipid extracts of the wing imaginal disc. This data set was utilized as a high mass accuracy MS reference data set for tissue specific lipids (*13*). We identified 156 lipid species with two distinct extraction methods (**Fig. 6 a**; for details see Material and Methods section) containing most of the major lipid classes previously assigned (*1*, *2*). Of the 156 lipid species 5 were MG, 22 DG and 10 TG. Seven species were identified by overlap of ToF SIMS and LC-MS methods **(Fig. 6b, Table 2)**. It is very likely that these seven lipids are molecular ions of intact MAGs, DAGs or TAGs.

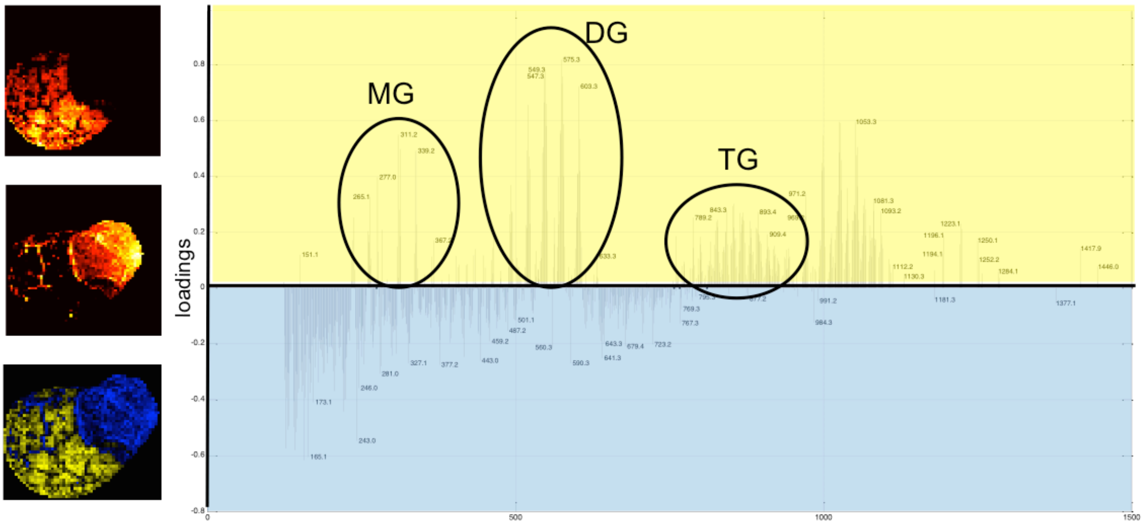


Fig. 5. Tentative assignments of peaks from ToF SIMS measurements: To assign the potential lipid species we used a recently published m/z catalogue with assigned lipids (*12*). We discovered that monoacyl-, diacyl- and triacylglyceroles are overrepresented in the wing blade part of the loadings spectra (blue spectral part of the loadings spectra).

The most abundant loadings spectra peak of the data sets acquired by the ToF SIMS in the positive ion mode (573.3m/z) was further validated by examination of the negative ion mode loadings spectra to decipher its fatty acid chain lengths. Clearly, the fatty acids C16:1 and C18:1 dominated the spectra for the component of the wing blade region **(Fig. 6c and d, Table 2)**. Furthermore, the lipid catalogue contains a DG identified as DG (34:2). Additionally, we used the LC-MS dataset as a reference dataset to check whether certain MAGs, DAGs, and TAGs are found by ToF SIMS but not being differentially distributed. From the ToF SIMS data we could identify five lipids found in the lipid catalogue **(Fig. S4)**. Taken together, the results presented here lead to an unambiguous identification of lipids distributed in domains delineated by genetically established compartment boundaries by ToF SIMS mass spectrometric imaging.

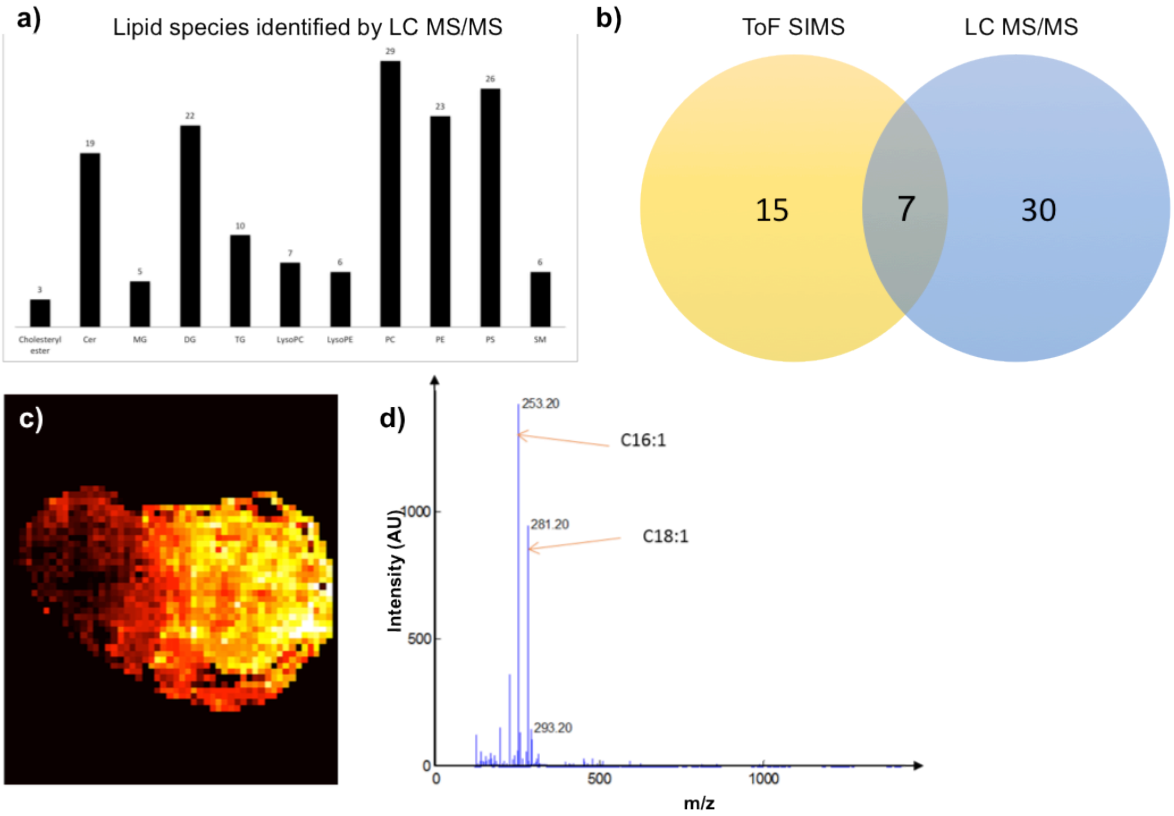


Fig. 6. Comparison between ToF SIMS annotation and LC-MS identified lipids: By lipid extraction and LC-MS analysis we identified 156 lipids in the wing disc (**a**). Thirty-seven of them are MAGs, DAGs or TAGs (A, red outlined box). In the ToF SIMS spectra we assigned totally 22 MG, DG, and TG species (**b**). Seven of these were also identified by the LC-MS approach (**b**, overlap). Strikingly, the most abundant peak of the ToF SIMS spectra (573.3m/z) is one of the seven lipids identified concurrently by ToF SIMS and LC-MS. The identity of this peak of the ToF SIMS spectra could further be verified using the negative ion mode pattern (**c**). The two most abundant fatty acids are C16:1 and C18:1 (**d**, red arrows). Cer = Ceramide, MG = monacylglycerol, DG = diacylglycerole, TG = triacylglycerole, PC = Phosphatidylcholine, PE = Phosphatidylethanolamine, PS = Phosphatidylserine, SM = Sphingomyelin

Lipid class (C:DB)	Passarelli et al. m/z	Adduct	ToF SIMS Rep.1 m/z	Rep.2 m/z	Rep. 3 m/z	Rep. 4 m/z	Rep. 5 m/z	Rep.6 m/z	LC MS/MS m/z	Adduct	Monoisotopic mass
MG(16:1)	311.3	(M+H-OH)+	311.2	311.1	311.1	311.1	311.1	311.1	n.d.	-	328.2614
MG(16:0)	313.3	(M+H-OH)+	313.2	313.1	313.1	n.d.	n.d.	313.1	348.3092	(M+NH4)+	330.277
MG(18:1)	339.3	(M+H-OH)+	339.2	339.2	339.2	339.2	339.2	339.2	n.d.	-	356.2927
DG(30:2)	519	(M+H-OH)+	519.3	519.3	519.3	519.3	519.3	519.3	n.d.	-	536.4441
DG(30:1)	521	(M+H-OH)+	521.3	521.3	512.2	521.3	521.2	521.2	n.d.	-	538.4597
DG(30:0)	523	(M+H-OH)+	523.3	523.3	523.3	523.3	523.3	523.3	n.d.	-	540.4754
DG(32:2)	547	(M+H-OH)+	547.3	547.3	547.2	547.3	547.2	547.2	n.d.	-	564.4754
DG(32:1)	549	(M+H-OH)+	575.3	575.3	575.3	575.3	575.3	575.3	567.4964	(M+H)+	566.491
DG(32:0)	551	(M+H-OH)+	551.3	551.3	551.3	551.3	551.3	551.3	568.5389	(M+NH4)+	568.5067
DG(34:3)	573	(M+H-OH)+	573.3	573.3	573.2	573.3	573.3	573.2	608.5209	(M+NH4)+	590.491
DG(34:2)	**575**	**(M+H-OH)+**	**575.3**	**575.3**	**575.3**	**575.3**	**575.3**	**575.3**	**593.511**	**(M+H)+**	**592.5067**
DG(34:1)	577	(M+H-OH)+	577.3	577.3	577.3	577.3	577.3	577.3	612.5544	(M+NH4)+	594.5223
DG(34:0)	579	(M+H-OH)+	579.3	579.3	579.2	n.d.	n.d.	579.2	n.d.	-	596.538
DG(36:4)	599	(M+H-OH)+	599.3	599.2	599.2	599.2	599.2	599.2	634.5365	(M+NH4)+	616.5067
DG(36:3)	601	(M+H-OH)+	601.3	601.3	601.2	601.2	601.2	601.2	n.d.	-	618.5223
DG(36:2)	603	(M+H-OH)+	603.3	603.3	603.3	603.3	603.3	603.2	n.d.	-	620.538
TG(50:3)	851	(M+Na)+	n.d.	n.d.	n.d.	n.d.	851.3	n.d.	n.d.	-	828.7207
TG(50:2)	853	(M+Na)+	853.4	n.d	853.3	853.3	853.3	853.3	n.d.	-	830.7363
TG(50:1)	855	(M+Na)+	855.4	n.d.	n.d.	n.d.	855.3	n.d.	n.d.	-	832.752
TG(52:3)	879	(M+Na)+	n.d.	n.d.	n.d.	n.d.	879.3	n.d.	n.d.	-	856.752
TG(52:2)	881	(M+Na)+	n.d.	n.d.	n.d.	n.d.	881.3	n.d.	n.d.	-	858.7676
TG(52:1)	883	(M+Na)+	n.d.	n.d.	n.d.	883.3	n.d.	n.d.	n.d.	-	860.7833

Table 2. List of identified MAGs, DAGs, and TAGs by ToF SIMS and LC-MS: The columns show the lipid class, mass spectra (m/z), adduct as well as the monoisotopic mass of each lipid identified. Yellow highlights point out lipids identified by both ToF SIMS and LC-MS approaches concurrently. The most abundant peak of the ToF SIMS loadings spectra also identified with LC-MS is written in blue letters**.**

Discussion

In this study, we applied ToF SIMS imaging mass spectrometry to analyze the distribution of small molecules in the wing imaginal disc of third instar *Drosophila melanogaster* larvae. We demonstrated that small molecules show distributions similar to known genetic patterns. Applying 3D coherent anti-Stokes Raman spectroscopy (CARS) we assessed the tissue topography and chemical integrity of the sample. Thus, we revealed the molecular nature of some peaks detected by ToF SIMS creating these patterns (*14*). For the B/W pattern, CARS constrained the ToF SIMS peaks to be lipids. However, orthogonal technologies need to be applied in order to unambiguously identify the molecules giving rise to these patterns in the ToF SIMS MSI. Mass isolation of imaginal discs in combination with FACS sorting of fluorescently labeled cells of different compartments followed by mass spectrometric investigation (e.g. proteomics, lipidomics) is one possible alternative method. For the B/W pattern we used published data resources to assign lipids to the m/z values from the B/W pattern found in positive ion mode. However whether the assigned lipids are from the molecular ion still needs to be verified. As ToF SIMS can lead to fragmentation of lipids, it is possible that the observed DAGs are fragments of intact TAGs or phospholipids. At the moment, this cannot be fully resolved as we find both intact lipids and DAGs in the LC-MS data with similar fatty acid chain length. But, in the ToF SIMS data sets in negative ion mode we do not see the head group (184m/z) of the possible intact lipid. Thus, it is unlikely that the peak of the DAGs is rather a fragment then the peak from the molecular ion.

To address a potential function of these lipids in pattern formation and maintenance it is crucial to understand their metabolic role and biochemical properties. Recent lipidomics studies revealed some surprising dynamics of DAGs and TAGs during larval development. It has been shown that DAGs and TAGs are quantitatively increasing in the wing imaginal disc during larval development reaching a peak at late third instar (*1*, *2*). TAGs are mainly considered storage lipids for energy via β-oxidation (*15*, *16*). However, TAGs are not used as an energy source via β-oxidation during puparation and metamorphosis (*1*). This strongly indicates that the TAGs have a different function in imaginal discs. In contrast to the TAGs the DAGs show different dynamics based on their fatty acid (FA) chain length. Medium FA chain DAGs drastically decrease at the transition from third larval instar (L3) to the pupa stage whereas the long chain DAGs are still quantitatively increasing. This indicates that the metabolism of different DAGs is important for distinct functions. Some DAGs serve as second messengers in several receptor signal transductions via the hydrolysis of membrane lipids into DAGs and inositol-3,4,5-triphosphate ($IP_{3)}$ by the enzyme phospholipase C (PLC). The generation of IP_3 leads to a release of Ca^{2+} from the endoplasmatic reticulum (ER). Ca^{2+} together with the DG signaling pathway governs processes like cell division and differentiation (*17*). Additionally, phospholipase D (PLD) activity is known to regulate cell growth and proliferation using Raf and mTOR as mediators (*18*). However, it remains to be determined what exact mechanism these DG pathways are controlling in Drosophila. Another way DAGs might contribute to compartment boundary formation and maintenance is by remodeling the membrane in response to effector signals. In the immune system for example successful phagocytosis requires actin remodeling by DG generation (*19*). Other mechanisms leading to similar membrane remodeling and polarization are in detail reviewed in (*20*). Moreover, it has been demonstrated that DAGs in combination with other membrane lipids in distinct proportions have an important function in determining the structure and dynamics of biological membranes (*21*). For example, the membrane composition of a cell is important for the curvature of the membrane itself on a single cell level (*22*). Likewise, if the membrane composition has an influence on the curvature of a single cell it is probable / presumable that the membrane composition of a sub-tissue region, as a compartment, may influence the tissue curvature. Such a phenomenon might be responsible for the generation of tissue folds, and our data seems to be consistent with this notion (**Fig.1** and **Fig.2**). Alternatively, an overrepresentation of DAGs in one cell population might increase the cells affinity to each other and thereby contribute to the segregation of cells of different populations. Moreover, we hypothesize that DAGs might also be differentially distributed in the two leaflets of the lipid bilayer as a consequence of the increasing tension generated during metamorphosis. At last, the intermediate long chain DAGs identified in the ToF SIMS datasets can also be used as building blocks for rapid synthesis of new phospholipids. As the organism undergoes tremendous change during pupation, it is likely that an enormous amount of membranes must be generated in a very short time. The fact that we identified lipids with the same fatty acid chain length in the LC-MS approach and the ToF SIMS datasets may support this theory. Nonetheless, it remains to be determined how different lipid compositions of compartments influence development and growth of imaginal discs.

In this study, we applied ToF SIMS imaging mass spectrometry to analyze the distribution of small molecules in the wing imaginal disc of third instar *Drosophila melanogaster* larvae. We demonstrated that small molecules show distributions similar to known genetic patterns. We submit that our results raise the important question of how small molecules are involved in developmental processes such as compartment boundary formation and maintenance. More importantly however, our workflow serves as a template for determining how small molecules can be simultaneously described in space and composition, a technique that closes a long-sought-after gap in experimental developmental biology.

References and Notes:

1. X. Guan *et al.*, Biochemical Membrane Lipidomics during Drosophila Development, *Dev. Cell* **24**, 98–111 (2013).

2. M. Carvalho *et al.*, Effects of diet and development on the Drosophila lipidome, *Mol. Syst. Biol.* **8**, 600 (2012).

3. E. Amstalden van Hove, D. Smith, R. Heeren, A concise review of mass spectrometry imaging, *J. Chromatogr. A* **1217**, 3946–4000 (2010).

4. A. M. Hicks, C. J. DeLong, M. J. Thomas, M. Samuel, Z. Cui, Unique molecular signatures of glycerophospholipid species in different rat tissues analyzed by tandem mass spectrometry., *Biochim. Biophys. Acta* **1761**, 1022–9 (2006).

5. B. Tyler, G. Rayal, D. Castner, Multivariate analysis strategies for processing ToF-SIMS images of biomaterials., *Biomaterials* **28**, 2412–2423 (2007).

6. F. Bonnier, H. Byrne, Understanding the molecular information contained in principal component analysis of vibrational spectra of biological systems, *Analyst* (2011).

7. G. Eijkel *et al.*, Correlating MALDI and SIMS imaging mass spectrometric datasets of biological tissue surfaces, *Surf. Interface Anal.* **41**, 675–1360 (2009).

8. E. Villa-Cuesta, E. Gonzalez-Perez, J. Modolell, Apposition of iroquois expressing and non-expressing cells leads to cell sorting and fold formation in the Drosophila imaginal wing disc, *BMC Dev. Biol.* **7**, 106 (2007).

9. I. Guillén *et al.*, The function of engrailed and the specification of Drosophila wing pattern., *Development* **121**, 3447–56 (1995).

10. Z. Movasaghi, S. Rehman, I. U. Rehman, Raman Spectroscopy of Biological Tissues, *Appl. Spectrosc. Rev.* **42**, 493–541 (2007).

11. S. Rangarajan, B. J. Tyler, Topography in secondary ion mass spectroscopy images, *J. Vac. Sci. Technol. A Vacuum, Surfaces, Film.* **24**, 1730–1736 (2006).

12. M. Passarelli, N. Winograd, Lipid imaging with time-of-flight secondary ion mass spectrometry (ToF-SIMS), *Biochim. Biophys. Acta* **1811**, 976–1066 (2011).

13. X. Gao *et al.*, A reversed-phase capillary ultra-performance liquid chromatography-mass spectrometry (UPLC-MS) method for comprehensive top-down/bottom-up lipid profiling, *Anal. Bioanal. Chem.* **402**, 2923–2933 (2012).

14. J. P. R. Day, G. Rago, K. F. Domke, K. P. Velikov, M. Bonn, Label-free imaging of lipophilic bioactive molecules during lipid digestion by multiplex coherent anti-Stokes Raman scattering microspectroscopy., *J. Am. Chem. Soc.* **132**, 8433–9 (2010).

15. R. A. Coleman, D. G. Mashek, Mammalian triacylglycerol metabolism: synthesis, lipolysis, and signaling., *Chem. Rev.* **111**, 6359–86 (2011).

16. K. Athenstaedt, G. Daum, The life cycle of neutral lipids: synthesis, storage and degradation., *Cell. Mol. Life Sci.* **63**, 1355–69 (2006).

17. T. K. Harden, J. Sondek, Regulation of phospholipase C isozymes by ras superfamily GTPases., *Annu. Rev. Pharmacol. Toxicol.* **46**, 355–79 (2006).

18. M. N. Hodgkin *et al.*, Diacylglycerols and phosphatidates: which molecular species are intracellular messengers?, *Trends Biochem. Sci.* **23**, 200–4 (1998).

19. C. C. Scott *et al.*, Phosphatidylinositol-4,5-bisphosphate hydrolysis directs actin remodeling during phagocytosis., *J. Cell Biol.* **169**, 139–49 (2005).

20. M. Almena, I. Mérida, Shaping up the membrane: diacylglycerol coordinates spatial orientation of signaling, *Trends Biochem. Sci.* **36**, 593–603 (2011).

21. H. Sprong, P. van der Sluijs, G. van Meer, How proteins move lipids and lipids move proteins., *Nat. Rev. Mol. Cell Biol.* **2**, 504–13 (2001).

22. I. R. Cooke, M. Deserno, Coupling between lipid shape and membrane curvature., *Biophys. J.* **91**, 487–95 (2006).

23. E. G. Bligh, W. J. Dyer, A RAPID METHOD OF TOTAL LIPID EXTRACTION AND PURIFICATION, *Can. J. Biochem. Physiol.* **37**, 911–917 (1959).

Acknowledgments: This project was in part financed with a grant from the Swiss SystemsX.ch initiative, grant IPP-2011/"Identification of Small Molecules Relevant for Patterning and Growth of the Wing Imaginal Disc Using High Resolution Mass Spectrometry" to EB, KB and RMAH. This work is part of the research program of the Stichting voor Fundamenteel Onderzoek der Materie (FOM), which is financially supported by the Nederlandse Organisatie voor Wetenschappelijk Onderzoek (NWO). We also thank the Kanton of Zurich and the Swiss National Science Foundation for support.

Supplementary Material:

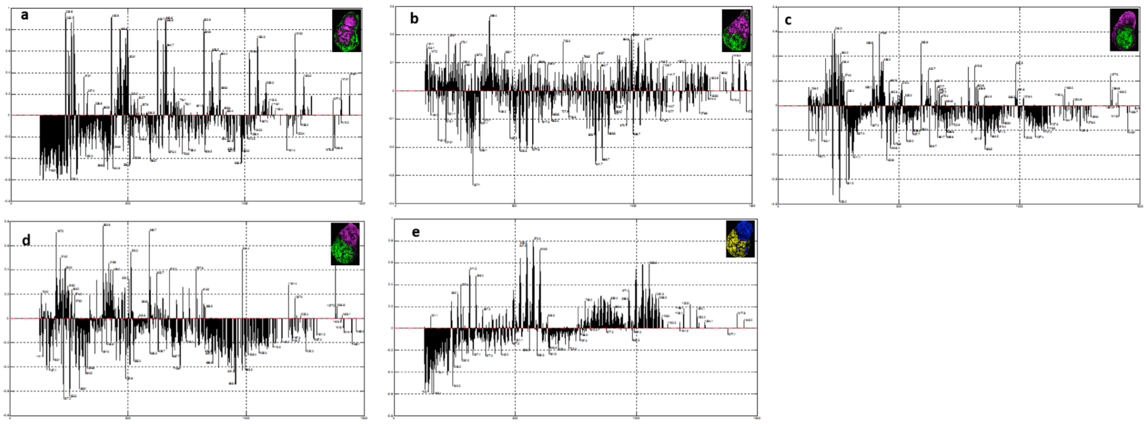


Fig. S1: Loadings spectra of different patterns observed by ToF SIMS. The spectra of the pattern A/P (**a**), D/V (**b**), pouch-non pouch (**c**) and the B/W in negative (**d**) and positive (**e**) ion mode respectively are depicted to assign the peaks to small molecules.

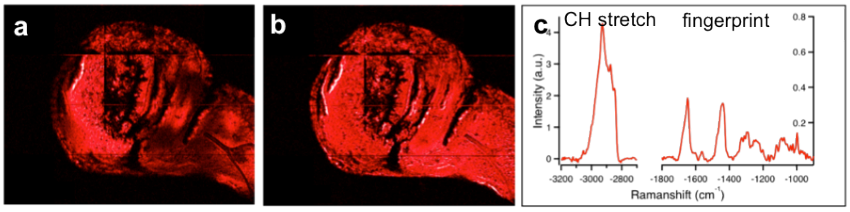


Fig. S2: Tissue integrity and 3D topography assessed with CARS. Two representative z-sections are depicted in (**a**) and (**b**). The CH-stretch and the fingerprint region in (**c**) are identical with the spectra obtained in literature.

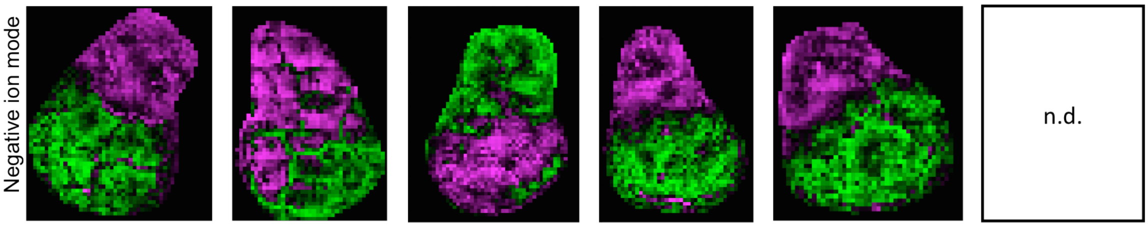
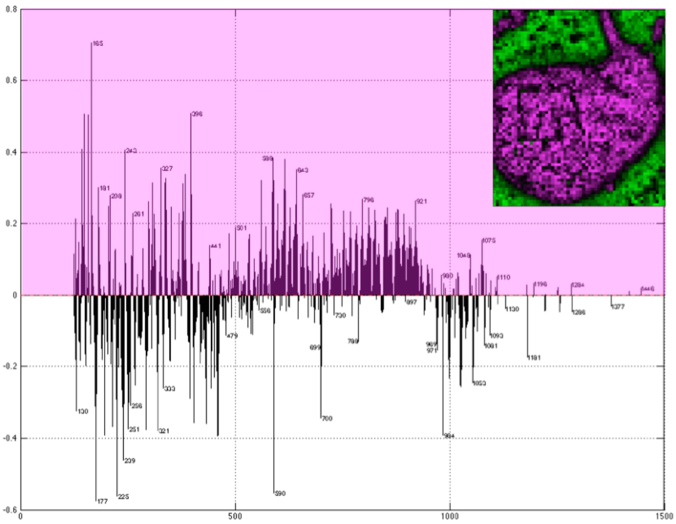


Fig. S3: Negative ion mode ToF SIMS PC output of the B/W pattern from the same wing imaginal discs as the repetitive measurements from **Fig. 4**. N.d. = not detected



ID C:DB	ToF SIMS m/z	M+H-OH m/z	M+Na+ m/z
MG(20:3)		363.3	403.3
MG(20:0)	369.1	369.3	409.3
MG(24:1)		423.4	463.4
DG(32:0)		551.5	591.5
DG(33:4)		557.5	597.5
DG(34:4)	571.2	571.5	611.5
DG(35:4)		585.5	625.5
DG(35:2)	589.7	589.5	629.5
DG(37:0)		621.6	661.6
DG(38:3)	629.3	629.6	669.6
DG(38:0)		635.6	675.6
DG(40:5)		653.6	693.6
DG(40:3)		657.6	697.6
DG(42:11)	669.3	669.5	709.5
DG(42:10)		671.5	711.5
DG(42:1)		689.6	729.6
DG(44:10)		699.5	739.5
DG(46:2)		743.7	783.7
DG(46:1)		745.7	785.7

Fig. S4: MAGs, DAGs, and TAGs identified by both ToF SIMS and LC-MS approaches concurrently which are not found to be differentially distributed in the B/W pattern.

Material and Methods

Sample preparation

For CARS wild type yellow white (yw) third instar larvae were used. For SIMS either yw strain? or genetically modified flies were used having the following genotype: yw UAS-mCD8::GFP hsp-flp; hh-Gal4/TM6b [26].

Wing imaginal discs were manually dissected in ice cold PBS and dehydrated in a 10x PBS solution (Sigma-Aldrich, Zwijndrecht, NL) for 30 min. For CARS and SIMS the discs were then mounted on conventional microscopy cover slides or, for SIMS only, they were mounted on conductive ITO slides. Mounted discs were washed with MS grade water (Sigma-Aldrich, Zwijndrecht, NL) three times to remove additional salts. For CARS the discs were directly analyzed without further sample preparation. For SIMS the discs were covered with a 2nm gold layer using a sputter coater (Quorums Technologies SC7640, Newhaven, UK) equipped with a FT7607 quartz crystal microbalance stage and a FT7690 film thickness monitor.

Microscopy

Discs were imaged with a Leica DMRX (Leica, Wetzlar, Germany) microscope equipped with an OSRAM HBO 50W/L2 short arc mercury lamp (Osram AG, München, Germany) at 10x magnification.

Data acquisition

SIMS MSI:

SIMS was performed a TRIple Focusing Time-of-flight ion microscope (TRIFT-II), a Physical Electronics (PHI) instrument. The TRIFT-II mass spectrometer is equipped with a gold liquid metal ion gun (LMIG). Twenty-two keV Au+ primary ions were micro focused on the sample surface. The total primary ion dose density amounted was kept well below the static limit. Positive or negative secondary ions were extracted to the mass analyzer with a static voltage of -/+ 3.5kV and post accelerated in front of the detector (dual stage MCP) by an additional 10kV. Signal from ions in the mass range 1-1500m/z was recorded. Full wing discs were imaged step-by-step in a mosaic formed by 8x8 individual tiles between which the stage moved. Each tile of about 80-95µm width was probed by the primary ion beam in a 256×256 raster for a duration of 30 sec to ensure that the total ion dose was below the static limit of SIMS. The resulting image was saved as a RAW file for further data processing.

CARS spectro microscopy:

A dual-output laser source (Leukos-CARS, Leukos, Limoges, France) provides the pump and Stokes beams. The source is a passively Q-switched 1064-nm microchip laser, delivering < 1-ns pulses at 32kHz repetition rate and ~300mW average power. The laser beam is equally divided into two parts with a beam splitter. One part is sent through a bandpass filter (FL1064-10, Thorlabs) and used directly as the pump beam. The other part is introduced into a photonic crystal fiber that creates super continuum emission of 420–2400nm at the fiber output, with > 100µW nm^{-1} spectral power density from 1.05µm to 1.6µm. The supercontinuum is coupled out of the fiber with a reflective collimator (RC04APC-P01, Thorlabs) and passed through 700-nm (FEL0700, Thorlabs) and 830-nm (LP02-830RS-25, Semrock) longpass filters. The Stokes and pump beams are overlapped at a dichroic mirror (LP02-1064RU-25, Semrock) and introduced into a modified inverted microscope (Eclipse Ti-U, Nikon). The pump and Stokes pulses are tightly focused onto the sample with a near IR objective (PE IR Plan Apo 100X, NA 0.75, Seiwa). The sample is mounted on nested stepper-motor-driven (Microstage, Mad City Labs) and piezo-driven stages (Nano-PDQ 375 HS, Mad City Labs) that together provide 25mm travel range with < 1-nm resolution. The CARS signal generated by the sample is collected in the forward direction by another objective (M-20X, NA 0.4, Newport) and sent through notch (NF03-532/1064E-25, Semrock) and short-pass filters (FES1000, Thorlabs) to remove the pump and Stokes beams. The filtered CARS beam is dispersed by a spectrometer (Shamrock 303i, 300 lines mm^{-1}, 1000-nm blaze, Andor) and detected on a deep-depletion CCD (Newton DU920P-BR-DD, Andor). The sample is raster scanned across the focal volume with steps of 1µm in plane and 2µm axially. Due to the large dimension of the wing disc, large three-dimensional images are reconstructed from adjacent tiles with in-plane dimension of 75x75µm or 50x50µm (specified for each measurement in the manuscript) and axial dimension determined by the thickness of the wing disc. For each position in the sample, a CARS spectrum in the range between -3400 and -600cm^{-1} is acquired. CARS images were acquired with pixel dwell times of 50ms.

Data analysis

SIMS MSI data:

The TRIFT associated software WinCadence (v4.4 Physical Electronics, Chanhassan, USA) was used to pick peaks on the total average mass spectrum and produce manually mass selected ion images of the wing discs. RAW images were converted to datacubes with the MSTools software suite developed at AMOLF (http://www.maldi-msi.org, in imzML / Software Tools). The datacube format enables direct observation of mass selected images while browsing through m/z values, and permits fast region-of-interest (ROI) selection accompanied by comparison between ROI average mass spectra. For principle component analysis the Matlab based chemometricks software package was used as previously described. Briefly, raw files were first converted to 1x1 pixel into .mat file formats to generate an average spectrum with a mass resolution of 0.1Da for the entire wing disc. The average mass spectrum was peak picked and about 1000 peaks were stored for every dataset. The peak list together with the initial RAW file was reconverted with 64x64 pixel spatial resolution for principle component analysis. Output of the TRIFTtricks software was manually investigated for patterns reminiscent of underlying biological knowledge. Patterns classified as interesting were displayed in false color images.

CARS data:

The quantitative information contained in the vibrational resonant term of the CARS spectra is hidden by the convolution with a non-resonant term arising from the electronic response of the material. The imaginary part of the third-order nonlinear susceptibility, Im[c(3)], is retrieved from the raw CARS spectra by the maximum entropy method (MEM), as described elsewhere [22]. Im[c(3)] spectra, referred to as retrieved CARS spectra , are directly comparable to spontaneous Raman spectra. Thus, the peak amplitude or integrated intensities are proportional to the number of scatterers in the focal volume.

Principal Component Analysis

PCA is a mathematical technique that describes the maximum amount of variance in a minimal amount of functions, the PCs. These PCs are linear combinations of the original variables (in this case m/z channels for SIMS and wave numbers for CARS). The correlation between each of the original variables and the calculated PCs is given by the PC loading. A plotted version of all loadings for a given PC is often used as graphical representation of that PC, the loading spectrum. By calculating the projection of each of the original spectra on the PC we obtain PC scores. The scores are used for creating a PC score image. PCA typically reduces the number of variables in a data set to a limited number of PCs that contain the major part of the variance in the data. The calculated PCs are ranked based on their total length, their Eigenvalue (EV). The number of relevant PCs is determined by using methods as the removal of PCs with EVs less than 1[32], using either a Scree plot [33] or choosing the n highest ranked PCs that account for a certain percentage of the total variance. PCA is a very efficient data compression technique as well as a noise reduction technique since noise in the data will be uncorrelated. Uncorrelated signals contain very little variance and thus will be described in the lower ranked PCs. Removal of these lower ranked PCs using one of the methods above-mentioned will result in a smaller reconstructed dataset containing less variables and less noise.

Canonical Correlation Analysis

CCA is a mathematical technique that determines the relationship between two separate sets of variables. The PC score matrices from both datasets are used to calculate their cross-covariance matrix. A PCA is performed on this cross-covariance matrix. This results in a new set of independent linear combinations of the original variables (in this case being PC scores), the Canonical Variates. The Canonical Variates represent the common data space for both datasets. Thus, we can calculate the scores for both datasets in the same data space and directly compare the scores by calculating the correlation. Spatial binning of the data was necessary to increase the signal to noise ratio and to reduce computer memory requirements.

SIMS

The SIMS PHI RAW data format was converted to Matlab/ChemomeTricks format using the in-house built sims2tricks software. Peak picking was performed on the SIMS data using the PEAPI algorithm.

CARS

The CARS data were exported from the analysis software Igor Pro as text data. The text data was converted to Matlab/ChemomeTricks data format using the in-house developed CARS2Tricks software. The separate CARS measurement tiles where stitched using the in-house developed StitchCARS software.

Data depth alignment

Static SIMS measurements image the topmost layer of the sample, independent of slight differences in height at the sample surface. Confocal CARS measurements, however, image a focal plane independent of differences in surface height (**Fig. 1**). In order to create data sets that reflect information from similar height regions as precise as possible, five CARS depth slices were used. Of each subsequent CARS layer starting at 0mm with a 2mm step size, the spectra with a summed intensity significantly higher than the background signal where stored. Combination of these data from the deepest level up to the highest level produced the CARS dataset that was used for the PCA and CCA analysis.

Data spatial alignment

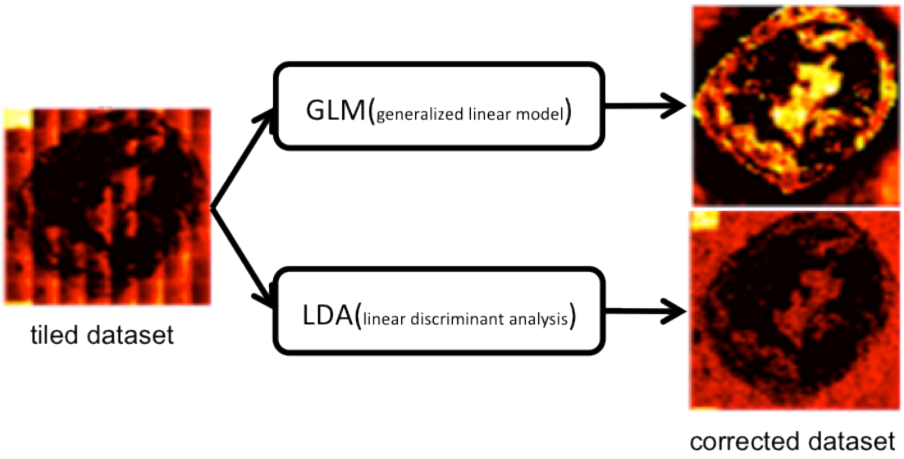
To be able to perform CCA on the data, these data must be spatially aligned as optimal as possible. Therefore, the SIMS data were translated, rotated, and scaled using MATLAB tools to obtain maximum overlap with the CARS dataset. The final data-preprocessing step involved the selection of the on-tissue pixel based on the total intensity obtained by the CARS measurement. All signals originating outside the tissue were removed from the data sets.

Lipid analysis by HPLC-electrospray ionization tandem mass spectrometry (HPLC-ESI MS/MS)

Third instar wing imaginal disc were collected using the mass isolation approach developed in house (unpublished data). Around 1000 individual discs were used per replicate. Lipids were extracted using modified Bligh and Dyer method (*23*). The extracts were removed, dried in vacuo and reconstituted in isopropanol before HPLC-electrospray ionization tandem mass spectrometry (HPLC-ESI-MS/MS). HPLC-ESI-MS/MS analyses were conducted on a Thermo Fisher Q Exactive mass spectrometer (San Jose, CA) with on-line separation using a Thermo Fisher/Dionex RSLC nano HPLC. HPLC conditions were: Atlantis dC18, 3 µm, 300 µm x 150 mm column (Waters Corporation, Massachusetts); mobile phase A, acetonitrile/water (40:60) containing 10 mM ammonium acetate; mobile phase B, acetonitrile/isopropanol (10:90) containing 10 mM ammonium acetate; flow rate, 6 µl/min; gradient, 10 % B to 60 % B over 5 min, 60 % B to 99 % B over 35 min and held at 99 % B for 10 min. Data-dependent tandem-MS scans were performed using one full MS scan [m/z 200 – 2000; 70,000 resolution (m/z 300)] followed by fragmentation in the HCD collision cell of the six most abundant ions in the precursor scan using a normalized collision energy (NCE) of 35 arbitrary units and mass analysis in the orbitrap at 17,500 resolution. Separate analyses were conducted using positive and negative ions detections.

Progenesis CoMet (Nonlinear Dynamics Limited, Newcastle, UK) was used to process the raw data files. Peak alignment and integration was performed and the relative abundance was generated. Lipid species were identified by searching the following databases: METLIN (http://metlin.scripps.edu/index.php); lipid maps (http://www.lipidmaps.org/data/structure/); HMDB (Human Metabolome Database; http://www.hmdb.ca/) using a 5-ppm mass tolerance. The putative lipid identifications were manually verified through examination of the tandem mass spectra and in comparison with the retention times from commercially available standards (Avanti Polar Lipids, Alabaster, AL). The confirmed lipids species are reported in this study.

Multi order correction algorithms to remove image distortions from mass spectrometry imaging datasets



analytical chemistry

Article

pubs.acs.org/ac

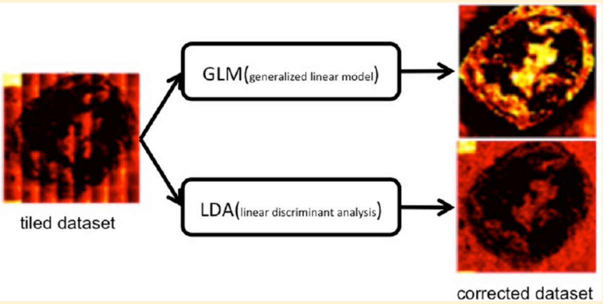
Multiorder Correction Algorithms to Remove Image Distortions from Mass Spectrometry Imaging Data Sets

Florian Gerber,[†,∥] Florian Marty,[‡,§,∥] Gert B. Eijkel,[§,∥] Konrad Basler,[‡] Erich Brunner,[‡] Reinhard Furrer,[*,†] and Ron M. A. Heeren[*,§]

[†]Institute of Mathematics, University of Zurich, Winterthurerstrasse 190, CH-8057 Zurich, Switzerland
[‡]Institute of Molecular Life Sciences, University of Zurich, Winterthurerstrasse 190, CH-8057 Zurich, Switzerland
[§]FOM-AMOLF, Science Park 104, 1098XG Amsterdam, The Netherlands

Supporting Information

ABSTRACT: Time-of-flight secondary ion mass spectrometry imaging is a rapidly evolving technology. Its main application is the study of the distribution of small molecules on biological tissues. The sequential image acquisition process remains susceptible to measurement distortions that can render imaging data less analytically useful. Most of these artifacts show a repetitive nature from tile to tile. Here we statistically describe these distortions and derive two different algorithms to correct them. Both a generalized linear model approach and the linear discriminant analysis approach are able to increase image quality for negative and positive ion mode data sets. Additionally, performing simulation studies with repetitive and nonrepetitive tiling error we show that both algorithms are only removing repetitive distortions. It is further shown that the spectral component of the data set is not altered by the use of these correction methods. Both algorithms presented in this work greatly increase the image quality and improve the analytical usefulness of distorted images dramatically.



Imaging mass spectrometry is a technology with increasing use in the bioanalytical field. The capability to obtain molecular images and at the same time chemical identity makes it a great tool to analyze a wide variety of organic and inorganic samples.[1] Secondary ion mass spectrometry imaging (SIMS) is one of the most used techniques to generate ion images of different molecular species. In a typical SIMS imaging experiment a primary ion beam is rapidly rastered over the sample of interest in a predefined track. At each position the primary ion beam is rastered over a squared surfaces area of a defined size that contains a fixed number of pixels (256×256). This is referred to as a tile. By decreasing the tile size the researcher can increase the spatial resolution down to the submicrometer level. A larger area is analyzed by performing a mosaic of a large number of adjoining tiles (e.g., 8×8).[2]

In a typical SIMS experiment the surface is bombarded with primary ions with energies in the kiloelectronvolt range. As a consequence, secondary ions are generated from the analyzed surface. Generally the ionization efficiency during this process is very low and decreases toward higher m/z values.[2] The secondary ions are analyzed in a mass spectrometer to generate position-specfic mass spectra. After completing the measurement at every (x, y)-position (tile) the ion images can be reconstructed. These images facilitate the investigation of the distributions of elements and small molecules such as lipids and other metabolites.[3−5]

For data interpretation it is crucial that the images have a very high contrast, as well as sharp borders between sample and sample holder. Yet, concealed measurement-based artifacts may vitiate the data set and subsequently impair the analysis. These distortions predominantly occur in every tile and show a high reproducibility from tile to tile. We therefore call these reproducible image errors tiling errors.

The cause of these errors is not completely understood. One of the possible reasons arises from the physical nature of the operating principle of SIMS and the surfaces analyzed. Most biological surfaces are considered electrical insulators.[6] During the process of bombarding the insulator with primary ions the sample will start to charge up. This charge up will influence or even prevent the emission of secondary ions.

Different methods have been developed to avoid these charge-up effects, for example, coating the surface of interest with conductive materials (e.g., gold nanolayer) or charge compensation by electron bombardment. These applications tremendously improved image quality in SIMS imaging mass spectrometry.[7]

But, charge buildup is not the only source of tiling errors. Misalignment of the primary ion gun or local detector sensitivity issues can also lead to similar distorted images. Unfortunately, most of these errors are only visible after the entire image has been acquired. Therefore, it is of great importance to generate

Received: July 3, 2013
Accepted: October 4, 2013
Published: October 4, 2013

dx.doi.org/10.1021/ac402018e | *Anal. Chem.* 2013, 85, 10249−10254

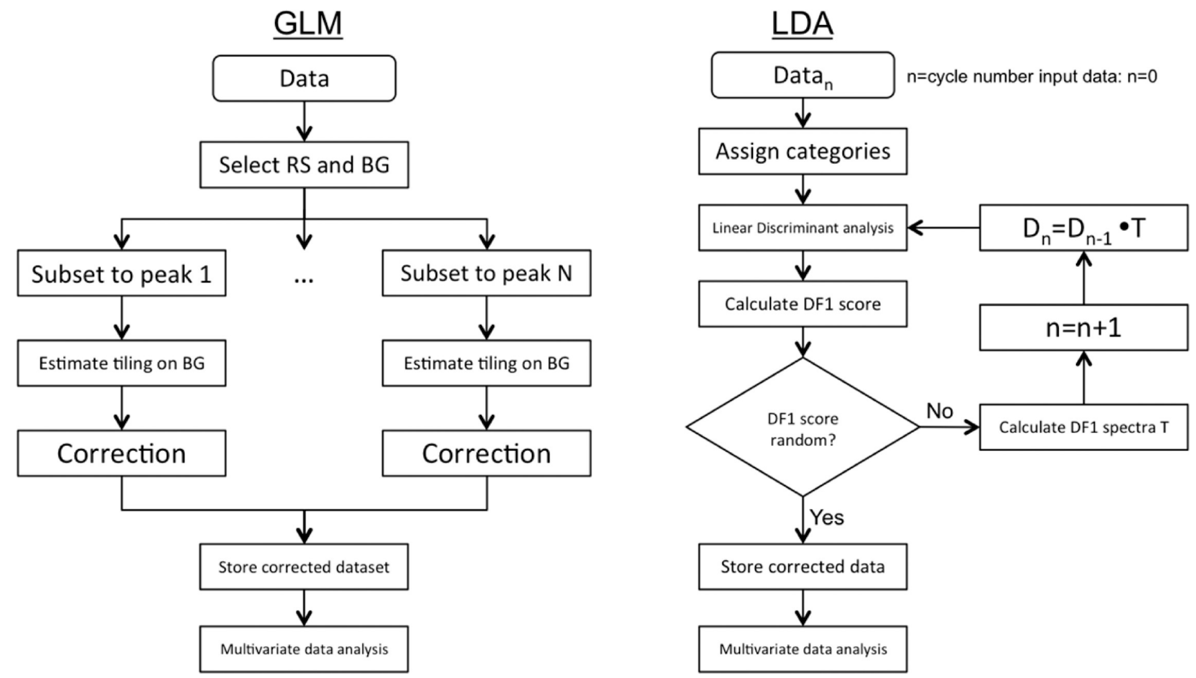


Figure 1. Workflow of the GLM and LDA approach used in this work. The GLM approach corrects the tiling error in a peak-by-peak manner, whereas the LDA performs linear discriminant analysis on the entire spectra of an assigned category. Due to its noniterative nature the GLM approach does not require any statistical test to stop the correction procedure. In contrast, LDA is iterative and hence has to stop when the DF_1 score calculated is random. The randomness is tested using a Wald−Wolfowitz runs test (ref 15).

tools that can correct for these tiling artifacts after the data acquisition.

Because of the repetitive nature of the tiling errors it is likely that these errors can be stochastically described. Such a stochastic model of tiling error would possibly allow the construction of methods to remove these artifacts. As a direct benefit the image quality would be enhanced, and therefore, multivariate analysis methods such as principal component analysis (PCA) may gain significance. In this paper we show that the tiling errors can indeed be statistically explained based on spectral features, and we have developed two algorithms that can directly remove these artifacts efficiently on the spectral level of the data. The algorithms and analytical approach described in this paper allow the user to get the most out of data acquired with tiling errors and thereby improve the data interpretation.

■ EXPERIMENTAL SECTION

Data Acquisition. Data used in this study was from atherosclerotic segments of mice arteria that were analyzed for lipid distribution. Additional information about the data sets can be found in ref 8. Briefly, the time-of-light (TOF) SIMS data were acquired on a Physical Electronics (Eden Prairie, MN) TRIFTII secondary ion mass spectrometer. Data sets were acquired in positive and negative ion mode ensuring that the total ion dose was kept below 10^{13} ions/cm^2, the so-called static limit of SIMS. Data acquisition was performed using the software WinCadence, version 4.4.0.17 (Physical Electronics, Chanhassan, MN). Data analysis was performed using the ChemomeTricks software package.[9]

Generalized Linear Model Approach. In this approach the measurement region of the detector is divided into two subregions, namely, the region of the sample (RS) as the region containing the measured tissue and the background (BG) as the region where we observe sample holder (glass) only. First, one single m/z-peak is corrected. To correct an entire data set we repeat the procedure for each peak (Figure 1). We assume a parametric description of the distortion and estimate its parameters on the BG where the tiling error pattern is not disturbed by biological information. The tiling error is then extrapolated to the RS, and the detector signal is separated into a tiling error and biological information component. The resulting corrected data is used as input for further multivariate data analysis (Figure 1).

The approach is based on the following two assumptions on the data structure: (A1) the tiling error in each peak can be described by a common model formulation; (A2) the tiling error structure on the RS is the same as on the BG. The first assumption allows us to perform the correction from peak to peak without altering the model.

We denote the observed intensity count at a pixel/location (x, y) with $z(x, y)$. For each m/z count we assume a Poisson distribution with rate $\lambda(x, y)$ and model the log rate $\log(\lambda(x, y))$ additively through a biological information part $\text{bioInfo}(x, y)$, a tiling error part $\text{tiling}(x, y)$, a constant intercept on the RS μ_{RS}, and an unstructured noise component ε. On the RS we have

$$\log(\lambda_{\text{RS}}(x, y)) = \mu_{\text{RS}} + \text{tiling}(x, y) + \text{bioInfo}(x, y) + \varepsilon \tag{1}$$

Assumption A2 implies that the tiling error is common over the entire slide, and for the BG the log rate is thus

$$\log(\lambda_{\text{BG}}(x, y)) = \mu_{\text{BG}} + \text{tiling}(x, y) + \varepsilon \tag{2}$$

as there is no biological information but a potentially different constant intercept μ_{BG}. To model the tiling error on the BG a generalized linear model (GLM) with log-link function is used.[10] Its linear predictors have to capture the tiling error. More specifically, we use fourth-order polynomials in terms of $(x \bmod k)$ and $(y \bmod k)$, with k the size of an individual tile. The modulus construction imposes a repeated structure on the model for the tiling error component.

Finally the estimated tiling error is removed from the data accordingly. The corrected rate linked to the biological information can be estimated as

$$\exp(\mathrm{bioInfo}(x, y)) = z_{\mathrm{RS}}(x, y)/\exp(\mu_{\mathrm{RS}} + \mathrm{tiling}(x, y)) \quad (3)$$

where $z_{\mathrm{RS}}(x, y)$ is the observed intensity count at a pixel/location (x, y) on the RS. $\exp(\mathrm{bioInfo}(x, y))$ does not have a tiling error component anymore.

The correction routine is implemented in the statistical software R.[11] The R package R.matlab is used to access the data from ChemomeTricks software package[9,12] and later to return data structures to the same software to perform PCA. Parallel computing is exploited with the R package multicore,[13] resulting in a virtually linear speed up. The first principal component of the data set is used to select the RS, and obvious outliers from the BG are excluded. In case the peak has a low overall intensity, the estimated tiling error may be very small for some (x, y), leading to unstable results. The fitting procedure is further improved by transforming the ratio in eq 3 to

$$(z_{\mathrm{RS}}(x, y) + q)/(\mu_{\mathrm{RS}} + \exp(\mathrm{tiling}(x, y)) + q) \quad (4)$$

where $q = \max(0, 1 - \exp(\min(\mu_{\mathrm{RS}} + \mathrm{tiling}(x, y)))$. The corrected intensities are rescaled to have the same overall intensity as the original ones, allowing comparisons of total ion images between different methods.

Linear Discriminant Analysis Approach. The second correction method assumes the correlation of the structure of the data with its position of the pixel in each tile. We used a linear discriminant analysis (LDA) approach to model and remove the tiling error from the data.[14]

The first discriminant function (DF_1) is defined as

$$\frac{D_1^{\mathrm{T}} B D_1}{D_1^{\mathrm{T}} W D_1} = \text{maximal}$$

where D_1^{T} is the transposed DF and B and W are the between and within group covariance matrices. The same definition holds for the second and higher DFs with the precondition that all DFs are orthogonal.

This approach first assigned a category (i.e., class) to each pixel according to their position within each tile. Then LDA was used to create a model based on these classes. The resulting first discriminant function (DF_1) described the common behavior of the spectra at each tile position over all tiles and therefore represented the tiling error. DF_1 scores of all pixels were back-transformed to the original (mass) feature space by multiplication of the DF_1 score with DF_1 loadings, and autoscaling and normalization were reversed. The resulting data set was subtracted from the original data set and the LDA procedure was repeated until no more tiling error related variance remained in the data. The procedure is summarized in the following equation:

$$S_{\mathrm{corr}} = ((S - \mathrm{ScoreDF}_1 \times \mathrm{LoadDF}_1) \circ \mathrm{Var} + \mathrm{Mean}) \times \mathrm{Norm}_s$$

where S_{corr} is the corrected spectrum, S is the autoscaled and normalized spectrum to be corrected, $\mathrm{ScoreDF}_1$ is the score of S, on DF_1. LoadDF_1 is the loading vector of DF_1, Var is the vector that contains the variances of the mass channels, Mean is the vector that contains the mean values of the mass channels, and Norm_s is the normalization factor for S. Please note that operator ∘ means the element-wise multiplication of the vectors.

After all the tiling error has been removed, the DF_1 scores would be random and are assessed with the Wald–Wolfowitz runs test on the DF_1 scores. If the Wald–Wolfowitz runs test is positive the cycling would be stopped.[15] The resulting corrected data were used as input for further multivariate data analysis (Figure 1).

Both presented algorithms in this study were tested on samples containing large amount of tiling distortions.

Simulation Study. To further support that the correction routines only remove tiling errors and do not alter the biological information, i.e., to correctly separate signal and repetitive noise, we have performed multiple simulation studies. Two representative examples are presented and discussed here. We took a corrected data set after the GLM correction was applied. This corrected data set represents the "ground truth" signal and was subsequently altered by adding different artificial tiling noise. The tiling pattern was regular or alternated by inverting every second tile (checkerboard-like) to simulate nonrepetitive noise. In the latter case, the correction routines should not detect a pattern and leave the data unaltered.

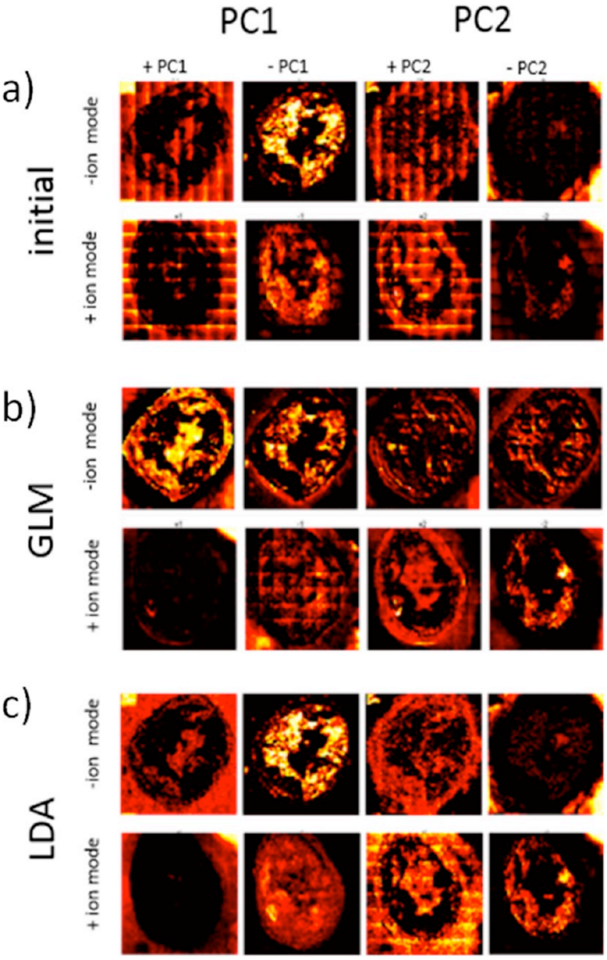


Figure 2. First and second principal components (PC1, PC2) of uncorrected data show strong image distortions both in positive and negative ion mode (a). Correction of the distorted data called initial with the GLM and LDA approach removed the distortions completely in negative ion mode (b and c). In positive ion mode some distortions are not removed. The remaining distortion needs to be further examined to improve the correction of it (b and c).

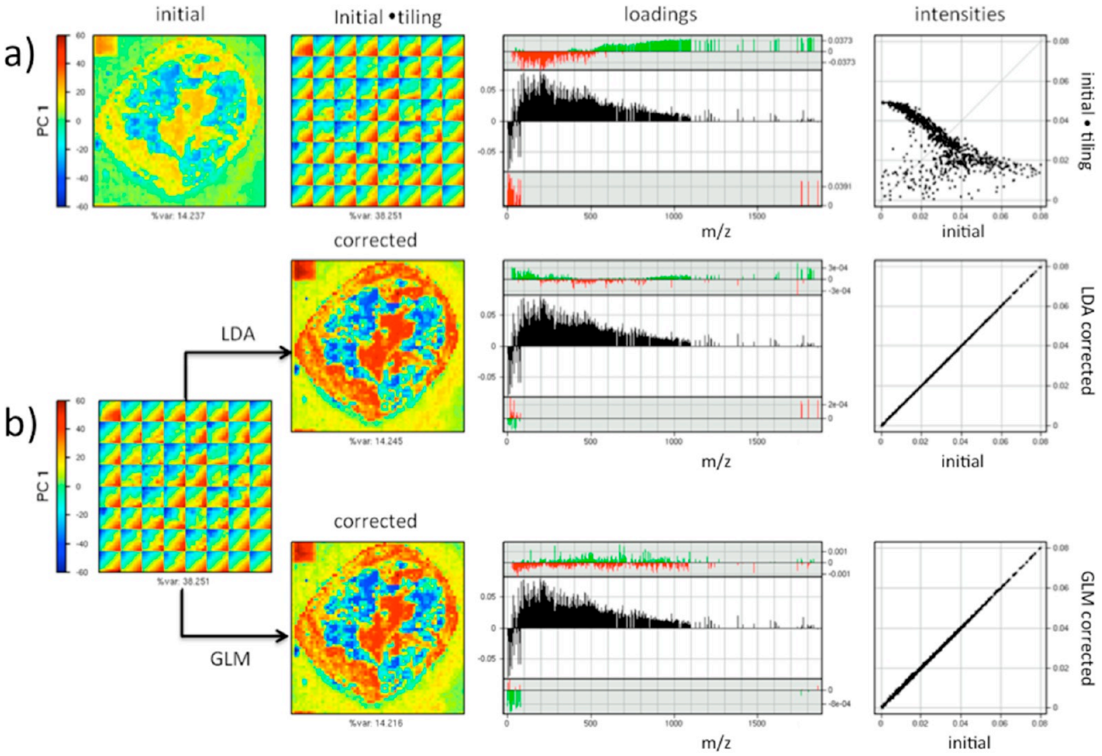


Figure 3. Corrected data set (initial) was distorted with a simulated repetitive tiling error pattern (initial·tiling) (a). This changed the loadings spectra and led to a significant shift of the peak intensities. As indicated above the loadings spectra, most of the loadings increase (green) in the higher mass range, whereas the lower masses decrease in loadings value by a factor of ±0.0379 (red). For the negative part of the principal component only minor changes of a factor of 0.0391 are observed as shown under the spectra indicated with green and red bars. The distorted simulated data set was then corrected with GLM and LDA (b). A direct comparison of the initial and the corrected data set does not reveal any changes in the loadings ($\pm 3 \times 10^{-4}$ for positive and $\pm 2 \times 10^{-4}$ for negative loading components) for the GLM. For the LDA approach the change is also negligibly small (±0.001 for the positive and $\pm 8 \times 10^{-4}$ for the negative loadings). In addition, for both algorithms the peak intensities remain the same after correction compared to the initial.

■ RESULTS AND DISCUSSION

Correction of Biological Data Sets. We developed two algorithms, GLM and LDA, to remove artificial repetitive distortions (tiling errors) (Figure 2a). The data sets used in here were arterial cross sections of atherosclerotic origin.[8] First, we corrected the data using GLM that is a parametric approach. The GLM approach uses the background (in our case, the sample holder) to estimate the tiling error structure and corrects it for one peak at the time. The procedure is then repeated for every peak (details, see the Experimental Section). Using appropriate model functions for the background, the biological information and the tiling error part, the tiled data can be corrected using a multiplicative correction approach. Subjectively, the algorithm performs well in negative ion mode and removed all the tiling error (Figure 2b). In positive ion mode compared to the noncorrected data set the algorithm removed most of the tiling error and facilitates image interpretation (Figure 2b).

Second, we applied a different correction algorithm based on LDA to correct the same data sets and compare for the suitability of the different algorithms. Comparable to the GLM approach the LDA approach removes the tiling in the negative ion mode (Figure 2c) and performs equally well in positive ion mode (Figure 2c). This indicates that the choice of an iteratively applied additive model seem to be as appropriate as the use of a multiplicative model. Both algorithms are not able to entirely remove the tiling error artifacts in the positive mode data set. Possibly, the remaining tiling error may not be as repetitive as assumed for the algorithms to be corrected. Further investigation of the remaining tiling errors are necessary to improve the correction efficiency for nonrepetitive artifacts.

Correction of Simulated Tiling Error on a Detiled Real-Life Sample. As both algorithms lead to a subjective improvement of the data quality we also examined their effect on the spectral components (principal component loadings). In a first step we added artificially generated regular tiling error to a corrected data set. The tiling error was constructed to resemble the typical tiling error patterns (Figure 3a). Clearly, the addition of tiling error heavily influences the loadings. This is indicated by the positive and negative differences between the loadings of the initial data set and the artificially tiled data set (red and green bars below and above the spectra Figure 3). But the GLM, as well as the LDA procedure, removed these artifacts completely (Figure 3b). Comparing the initial and corrected data set the loadings only vary slightly indicating that also the spectral distributions were restored to the initial state (again red and green bars below and above the spectra). This demonstrates that the algorithms do remove tiling error and thus increase the data quality. In an additional simulation we added artificially generated alternating tiling error which was not observed previously on real data (Figure 4a). We call this tiling error "checkerboard tiling". Again the loadings of the checkerboard-tiled

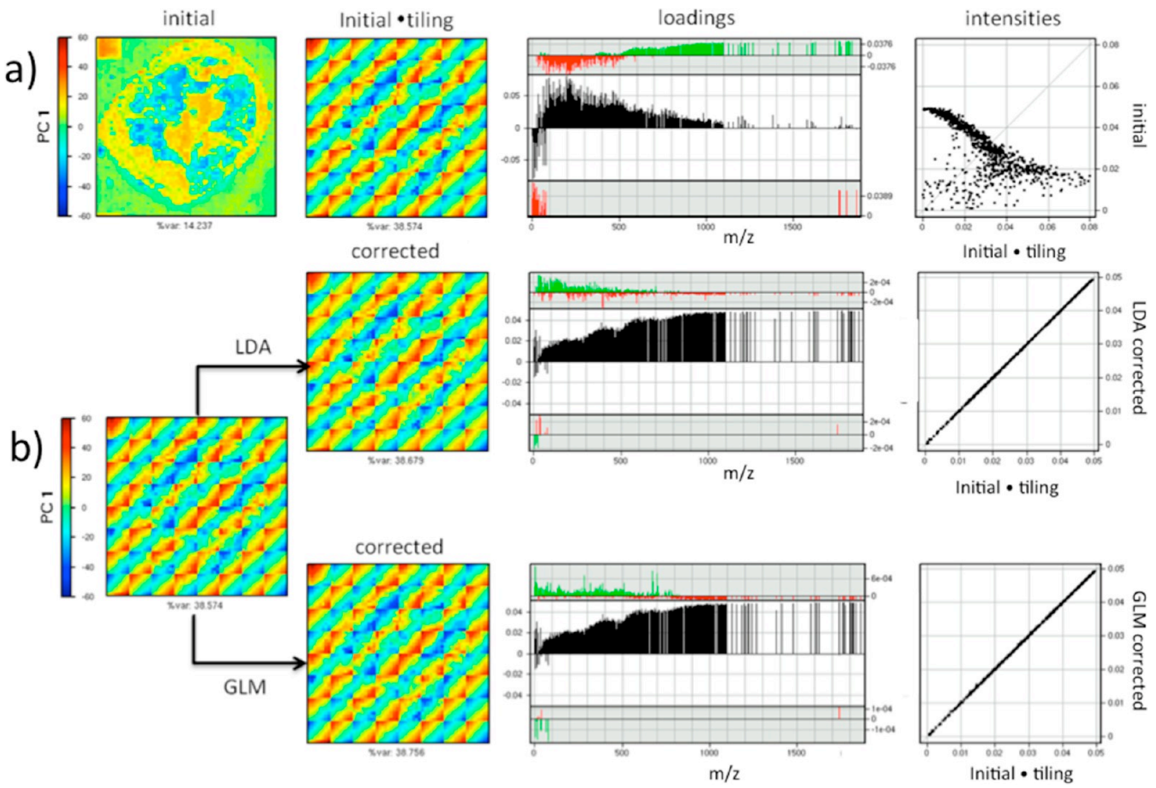


Figure 4. Corrected data set (initial) was distorted with a simulated nonrepetitive tiling error pattern (checkerboard tiling) (initial·tiling) (a). Comparable to the repetitive tiling error the loading spectra and the peak intensities are highly changed. As shown above the loadings spectra, most of the loadings increase (green) in the higher mass range, whereas the lower masses the loadings value decreases (red) by a factor of 0.0376. The negative part of the principal component does only change little (shown below, ± 0.0389). The distorted simulated data set was then corrected with GLM and LDA (b). As expected, both methods do not remove this distortion since it is not based on tile-by-tile repetition. Comparing the corrected data set to the initial·tiling data set the loadings and peak intensities remain unchanged ($\pm 2 \times 10^{-4}$ for the GLM and $\pm 6 \times 10^{-4}$ for the LDA positive loadings and $\pm 1 \times 10^{-4}$ for the negative loadings, respectively).

data set are different from the initial one (red and green bars below and above the loadings spectra Figure 4). Neither the GLM nor the LDA procedure removed this pattern (Figure 4b). Also the loadings were kept unchanged if the corrected data set was compared with the data set containing the distortions. (Note the small scales of the red and green bars.) This illustrates that the proposed algorithms only remove repetitive tiling error structures and no other potentially relevant information. Taken together these findings clearly demonstrate the use of such algorithms to increase image quality.

■ CONCLUSION

We have investigated a long-standing problem in mass spectrometry imaging, the occurrence of repetitive image artifacts called tiling error, through the use of two different algorithms. We demonstrated that a GLM and LDA approach are both equally well-suited to correct tiled data sets. Simulation studies of different tiling error patterns showed that the structure of the tiling error is only corrected if it is repetitive in every tile. Moreover, we demonstrated that both algorithms perform the correction without corrupting the spectral components of the data. Both algorithms have been applied to several dozen different samples. The versions described here represent a stable and comprehensive implementation. Both methods are built on different components that are almost orthogonal to each other (see Suppl. Table 1 in the Supporting Information). Naturally all these features can be interpreted as a disadvantage. For example, the GLM approach does not use the sample area to estimate the tiling error but subsequently uses a non-negligible area of background.

It is straightforward to build new and different correction approaches by exchanging or extending individual model components, e.g., using generalized additive models instead of GLM, using quadratic discrimination instead of a linear one. The utilization of the software R illustrates that there is no limitation to use proprietary software and that it is possible to build new or more complex correction routines. Finally, we believe that other fields with similar problems [e.g., matrix-assisted laser desorption/ionization mass spectrometry imaging (MALDI MSI), microscopy] may apply the algorithms presented herein, provided a repetitive error pattern is present.

■ ASSOCIATED CONTENT

Ⓢ Supporting Information

Additional information as noted in text. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Authors

*E-mail: heeren@amolf.nl
*E-mail: reinhard.furrer@math.uzh.ch.

Author Contributions

‖F.G., F.M., and G.B.E. contributed equally to the work.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This work is part of the research program of the Stichtingvoor-FundamenteelOnderzoek der Materie (FOM), which is financially supported by the NederlandseOrganisatievoor-WetenschappelijkOnderzoek (NWO). This project was in part financed with a grant from the Swiss SystemsX.ch initiative, Grant IPP-2011/ “Identification of Small Molecules Relevant for Patterning and Growth of the Wing Imaginal Disc Using High Resolution Mass Spectrometry” to E.B., K.B., and R.M.A.H. and by the University Research Priority Program (URPP) Systems Biology/Functional Genomics of the University of Zurich to R.F. The manuscript was written through contributions of all authors. All authors have given approval to the final version of the manuscript.

■ REFERENCES

(1) McDonnell, L.; Heeren, R. *Mass Spectrom. Rev.* **2007**, *26*, 606−643.

(2) Bertrand, P.; Weng, L. *Mikrochim. Acta, Suppl.* **1996**, *13*, 167−182.

(3) Debois, D.; Bralet, M.-P.; Le Naour, F.; Brunelle, A.; Laprevote, O. *Anal. Chem.* **2009**, *81*, 2823−2831.

(4) Cillero-Pastor, B.; Eijkel, G.; Kiss, A.; Blanco, F. J.; Heeren, R. M. A. *Anal. Chem.* **2012**, *84*, 8909−8916.

(5) Xu, K.; Proctor, A.; Hercules, D. M. *Mikrochim. Acta* **1996**, *122*, 1−15.

(6) Müller, G. *Appl. Phys.* **1976**, *10*, 317−324.

(7) Altelaar, A. F.; Klinkert, I.; Jalink, K.; De Lange, R. P.; Adan, R. A.; Heeren, R. M.; Piersma, S. R. *Anal. Chem.* **2006**, *78*, 734−742.

(8) Bot, M.; De Jager, S. C. A.; MacAleese, L.; Lagraauw, H. M.; Van Berkel, T. J. C.; Quax, P. H. A.; Kuiper, J.; Heeren, R. M. A.; Biessen, E. A. L.; Bot, I. *J. Lipid Res.* **2013**, *54*, 1265−1274.

(9) Eijkel, G.; Kokrer Kaleta, B.; Van Der Wiel, I.; Kros, J.; Luider, T.; Heeren, R. *Surf. Interface Anal.* **2009**, *41*, 675−685.

(10) Nelder, J. A.; Wedderburn, R. W. M. *J. R. Stat. Soc., Ser. A* **1972**, *135*, 370−384.

(11) R Core Team R: A language and environment for statistical computing. http://www.r-project.org.

(12) Bengtsson, H.; Riedy, J. *R.matlab: Read and write of MAT files together with R-to-Matlab connectivity*, 2012.

(13) Urbanek, S. *Multicore: Parallel processing of R code on machines with multiple cores or CPUs*, 2011.

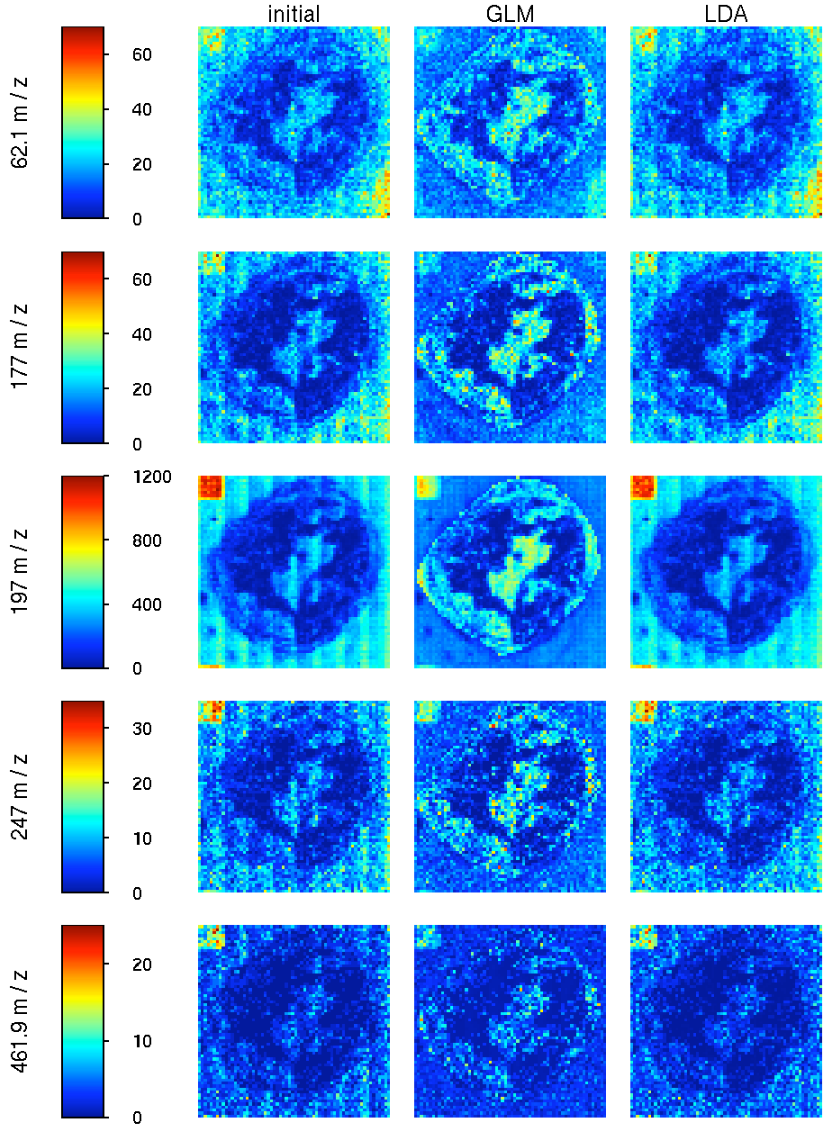
(14) Hoogerbrugge, R.; Willig, S. J.; Kistemaker, P. G. *Anal. Chem.* **1983**, *55*, 1710−1712.

(15) Bradley, J. V. *Distribution-Free Statistical Tests*; Prentice-Hall: Englewood Cliffs, NJ, 1968; p 388.

Supporting Information

GLM	LDA
Training/model tuning on background only	Using the entire slide
Underlying parametric model	Underlying non-parametric model
Multiplicative correction	Additive correction
Repetitive (over the peaks) procedure	Iterative (over the correction steps) procedure
No RS selection	RS selection
Use of open source software and ChemomeTricks	Entirely implemented in ChemomeTricks
Possibility to change model and correction parameters	Straightforward, out of the box functionality, no parameters needed

Suppl. Table 1. Summary of the correction features of the GLM and LDA, respectively.



Suppl. Figure 1. Images of five m/z peaks from the initial negative ion mode dataset (left column), their GLM correction (middle column) and their LDA correction (right column). The GLM images show a more pronounced tissue and a weakened background signal because of the distinction between tissue and background. The LDA approach does almost not change the tissue and background intensities. Some pixels in the LDA corrected images had slightly negative values which were set to zero. Both algorithms reduce the tiling of the images.

Discussion

To generate a fully functional organism the development of a multicellular structure needs to be tightly controlled and organized. Often, the development is organized around major body axes which control the main processes of development, such as patterning and growth. As often in biology, superior principles are also used on a smaller scale level. In terms of development this means that the body axes are also used to control development of organs. This phenomenon has been extensively studied in *Drosophila melanogaster.* Here, the classical boundaries A/P and D/V pattern the wing imaginal disc. Genetically the process of boundary formation is well understood (*7*, *10*, *12*, *14*, *61*). Additionally, the wing is further patterned along non-classical boundaries distinguishing between body wall and future wing blade or pouch vs. non-pouch tissue, respectively. Again, the genetics of these processes are very well understood (*22*, *23*, *62*). But, as stated previously, the role of the large pool of small molecules, such as lipids and carbohydrates, has not been investigated so far, mainly due to lack of technologies that allow the identification and localization of these molecular classes. Different reports have shown that MSI is able to chemically identify molecules directly on tissue. Additionally, the spatial information obtained during an MSI run is conserved making MSI a perfect tool to investigate the distribution of small molecules on the wing imaginal disc (*37*, *63–65*). Other potential techniques are also available and are extensively reviewed in (*35*, *59*, *66*). Here, we applied ToF SIMS imaging mass spectrometry to analyze the distribution of small molecules on the wing imaginal disc of third instar *Drosophila melanogaster* larvae. We favored SIMS MSI over MALDI MSI because of the limited spatial resolution available for MALDI at the beginning of the project (*35*). With a dimension of approximately 300x700µm, sub-micron resolution was needed to analyze the spatial distribution of small molecules. This can routinely be achieved with SIMS using Au^+ as a primary ion source (*43*). A major concern for ToF SIMS analysis is the topography of the tissue. It has been shown that topography can influence the spectra of ToF data tremendously and lead to misinterpretation of the achieved data. Several mechanisms have been developed and are reported to overcome this pitfall (*67–69*). In our study, we modified the sample protocol slightly to eliminate the three dimensional structure. The topography was assessed with CARS (*70*). A three dimensional reconstruction not only revealed no major topography remaining, but also showed that the tissue is still chemically intact considering the spectral investigation of the CARS spectra. Therefore, patterns observed in ToF SIMS analysis reflect the distribution of small molecules on the tissue rather than image artifacts.

At first, the ToF SIMS MSI data sets were analyzed in univariate mode with the Datacube explorer software (*52*). We did not observe a single peak being differentially distributed over the entire tissue. This corresponds with the previous observation that tissue subtypes can be distinguished rather on the profile of different molecules than compared to one single molecule (*71*). This may very well be the case for the compartments of the wing imaginal disc. Therefore, we analyzed the ToF SIMS data sets with a multivariate approach. We applied PCA to screen for PCs reminiscent of known genetic patterns. PCA has been extensively used in ToF SIMS and MSI field in general to analyze highly complex MSI data sets (*72–75*). We identified several PCs with distributions reminiscent of different classical and non-classical compartments. Among the classical compartments we observed the A/P and D/V compartmentalization. Of the non-classical compartments we identified the body-wall/wing blade (B/W) and the pouch/non-pouch compartment. Following, we correlated the patterns observed with fluorescent markers for A/P and morphological features for the B/W to rule out any bias in the interpretation of the PCs. Both, the A/P and the B/W pattern of the PCs, showed a high correlation with intrinsic fiducial markers. In order to get the molecular information it is crucial to investigate the loadings spectra of the PCs in both positive and negative ionization mode. Strikingly, only the B/W compartment could be observed in both negative and positive ion mode. The other compartments were only observed in negative ion mode. Further analysis of PC loadings spectra revealed a strong contribution of the gold peak for the A/P pattern. As the gold layer covering the wing imaginal disc is covered with is uniformly, an ion suppression like phenomenon apparently may give rise to the unequal distribution in the ToF SIMS image (*47*).

In the next step, we performed CARS spectro-microscopy to assess the molecular classes of the m/z values observed in the ToF SIMS data set. CARS only reproduced PCs reflecting the B/W pattern. With reference to literature we could assign a stronger protein signal to the body wall region. The wing blade region was richer in lipids. This fact massively narrowed the search space of the ToF SIMS data to find the molecules responsible for the B/W pattern. Therefore, we focused our follow up investigation on lipids only.

To further narrow down the molecular nature of the m/z values for the B/W pattern we investigated the loadings spectra of the PC from negative and positive ion mode. Applying a recently published resource of ToF SIMS peaks we could assign several lipid species to the wing blade region (*76*). Mainly diacyl- and triacylglycerol (DAG/TAG) lipids showed a differential distribution with a higher representation in the wing blade part. The fatty acid chain composition of the DAGs and TAGs could be assigned with the help of the negative ion mode spectra. However, whether the m/z values assigned to DAGs originate from the molecular ion still needs to be determined. As ToF SIMS can lead to heavy fragmentation it is possible that the observed DAGs are fragments of intact TAGs or phospholipids. At the moment, this cannot be fully resolved. But recent lipidomics studies revealed some surprising dynamics of DAGs and TAGs during larval development. It has been shown that DAG and TAG levels are increasing in the wing imaginal disc during larval development with TAGs reaching a peak at late third instar (*77*, *78*). Additionally TAGs are not used as an energy source via β-oxidation during pupation and metamorphosis (*77*). This strongly indicates that the DAGs and TAGs have a different function in the wing imaginal disc. However, these studies lack the spatial resolution to discuss the intra organ distribution and thus potential roles in development. Moreover, the peaks observed in the ToF SIMS spectra need to be unambiguously identified to find out more about their function. To this end, we created a wing imaginal disc specific lipid catalogue LC-MS/MS (*79*). We identified 156 different lipids from these extracts. Among the lipids identified 37 were DAGs and TAGs, respectively. Out of these 37 we found 9 in the loadings spectra of the corresponding patterns. This combinatorial approach of assignment and identification led to a high confidence considering the molecular nature of these peaks.

Additionally, high end imaging mass spectrometry tools could be used to further validate the peaks. Using C_{60} FT ICR and MALDI FT ICR with spot sizes of about 15µm we performed measurements to explore the potential of these new tools (*36*). Due to its high mass accuracy and resolving power the peaks could be identified using high mass accuracy MS1 identification (*80*, *81*). Preliminary data acquired on a C_{60} FT ICR show the potential this sort of instruments have. Especially the high mass accuracy will tremendously speed up the identification of small molecules. However, these instruments still need further improvement to be applied routinely.

Altogether, the workflow presented shows an innovative multidisciplinary way to analyze the distribution of different molecular classes on tissues. Moreover, part of the same workflow can also be used to investigate the potential role of the molecules observed. For example, the compartment boundaries can be shifted using the available genetics toolbox in Drosophila. The distribution of the small molecules can then be analyzed by ToF SIMS MSI and correlated with known compartment specific markers. If the small molecules are shifting together with the known markers it is likely that the underlying genetics do control the synthesis and localization of the small molecules.

Thus, lipid localization is controlled by the same mechanisms as pattern formation or is a direct product of these processes. However, if the distribution of the small molecules does not shift accordingly, the explanation is more complicated. Likely, in this case the distribution of the lipids is not related to compartmentalization at first. How the distribution in this case is achieved remains challenging to explain. Cell sorting experiments using labeled compartments (e.g. *hh*-Gal4;UAS-CD8-GFP) in combination with transcriptomics, proteomics and lipidomics can further be applied to gather more knowledge about the distribution of theses molecules. Moreover, this would also allow us to generate quantitative read outs of the distribution for lipids as done for proteins (Simigdala et al., unpublished data) and give new insights in the biology of boundaries.

Supporting Information

The following section was founded by collaboration with COST BM1104 "Mass Spectrometry Imaging: New Tools for Healthcare Research". The grant obtained was considered to be a short-term scientific mission (STSM) grant to work in one of the laboratories in the BM1104 network. The work was carried out in the lab of Prof. Dr. Ron M. A. Heeren. The following section is based on the Report submitted after completion of the STSM.

Using CyToF antibodies in SIMS imaging mass spectrometry

In CyToF, alternatively called mass cytometry, metal-labeled antibodies are used to indirectly measure different proteins (quantitatively) of interest. Most commonly in the CyToF field Ln are used as metal tags (*82*, *83*). Ln has the advantage that they are occurring with very low abundance in biological tissues. The fact that Ln has very good ionization properties in secondary ion mass spectrometers (SIMS) makes Ln-labeled antibodies promising candidates for MSI.

In order to test if Ln-labeled antibodies can be utilized in SIMS MSI we used Neodymium isotope 146- (Nd146-) labeled anti-mouse CD8 antibody directed against CD8-GFP expressed in the posterior compartment of the *Drosophila melanogaster* wing imaginal disc.

Prior to SIMS MSI experiments we tested extracellular staining of CD8-GFP with the same monoclonal antibody. We stained wing imaginal discs expressing a mouse CD8-GFP in the posterior compartment only with an anti-mouse CD8 antibody labeled with Nd136). SIMS MSI measurements were performed on a TRIFT II ToF SIMS mass spectrometer equipped with an Au^+ liquid metal ion gun (LMIG) as primary ion source. The sample was measured in positive and negative ion mode at a raster size of 100µm/tile. Each tile was measured for 30sec. With these settings the primary ion dose was well below the static limit of SIMS (10^{12}ions/cm2) to prevent fragmentation of biomolecules.

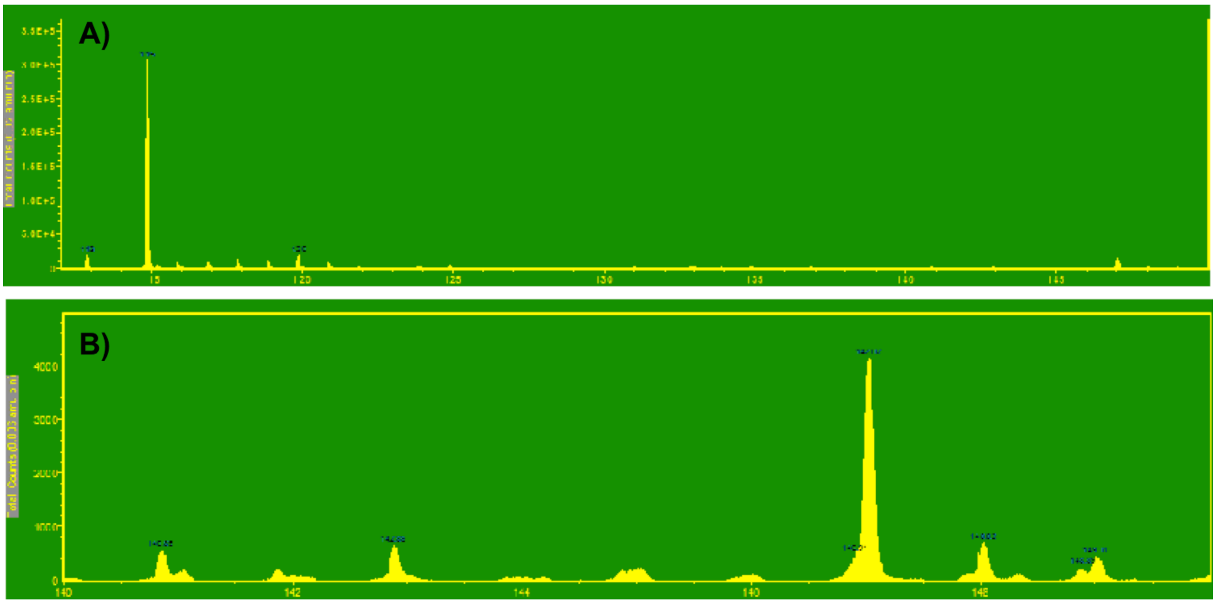


Figure 7: Representative spectra of the measured wing discs (positive mode)A) m/z range from 0-150 with dominant peak at 115m/z most likely the In^{+} from the ITO slide used to mount the sample onto **B)** Zoomed to 140-150m/Z of spectrum A). Strong peak at 147m/z which could be the NdH^{+} or more likely PDMS contamination

To the best of our knowledge Nd146-labeled antibodies have not been analyzed in SIMS so far. Therefore, we did not know which m/z values were to be expected. Most likely, Nd146 should show a strong peak at m/z=146 for the Nd^{+}. Additionally, the Nd146 may be double or triple charged leading to an m/z of 73 and 48.6, respectively. Further, Nd146 may also occur as an NdH^{+} with an m/z of 147. Considering the mass spectra we observed strong peaks at m/z=73 and m/z=147, but a very light signal at m/z=146 (Figure 1). As 73m/z and 147m/z are also related to Polydimethylsiloxan (PDMS) peaks, we checked the spectra for further PDMS related peaks. We detected a whole series of known PDMS peaks which showed no position-specific distribution. Since the antibodies used are highly specific, we would have expected any of these peaks to show up only in the posterior compartment of the imaginal disc. Unfortunately, none of the above mentioned peaks were detected specifically on tissue only (Figure 2A-D). Therefore, we conclude that the peaks observed at 73m/z and 146m/z are related to PDMS and not the Nd2^{+} or NdH$^{+,}$ respectively. To examine / see if Nd146 chelated to the antibody can be detected at all, we spotted pure antibody onto a target plate and analyzed the spot in positive and negative ion mode.

Again, we could not detect any signal at 146m/z (Figure3). Thus, we performed an imaging experiment on the same spot with 100µm tile size and 30min acquisition per tile. Therefore, we were well above the static limit of SIMS to fragment the antibody and the polymer to observe Nd. Only very little signal at 146m/z could be detected with again a big peak at 147m/z. This indicates that the PDMS contamination comes from the solution in which the antibody is dissolved. We did not further examine the exact origin of PDMS. However, we wanted to check if we could see any Nd in the SIMS MS. As DVScience produces de-isotope Ln, we acquired an Nd142 isotope together with a self-labeling kit in order to compare self-labeled antibodies to the commercially available pre-labeled antibodies. We measured all buffers and polymer they provide including the NdCl solution. Neither the buffers nor the polymer gave any signal which could be traced back on the imaginal discs measured previously. The 2µl Nd142 spot gave a strong signal at 141.87 of about 10,000 counts (Figure4). Applying only a tenth of the amounts used for coating of the antibody this signal intensity was likely to be a good indication that a Nd coupled antibody can be detected. To our surprise we saw a signal at 142.87m/z of about half the intensity of the 141.87m/z. This demonstrates that either the Nd142 was not 99.95% pure as described in literature or that we also see the NdH^+. Further studies need to be performed in order to rule out either possibility.

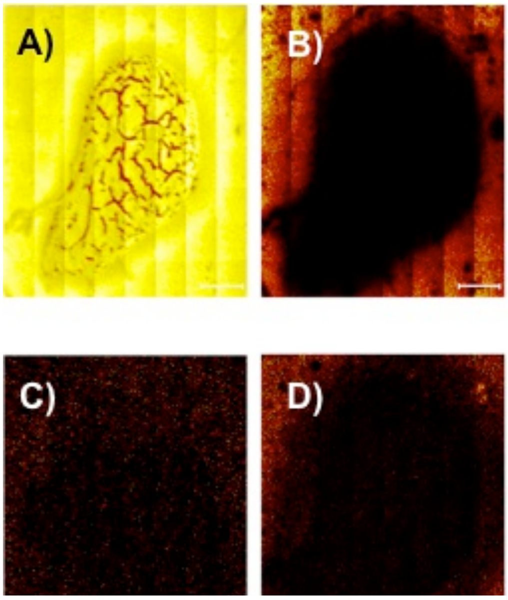


Figure 2: SIMS MSI of wing imaginal disc labeld with anti-mCD8-Nd146. A) Total ion image **B)** m/z 115 representing In^+ oft the ITO slide. **C)** m/z 146 of the Nd146 **D)** m/z 147 being either NdH^+ or PDMS related

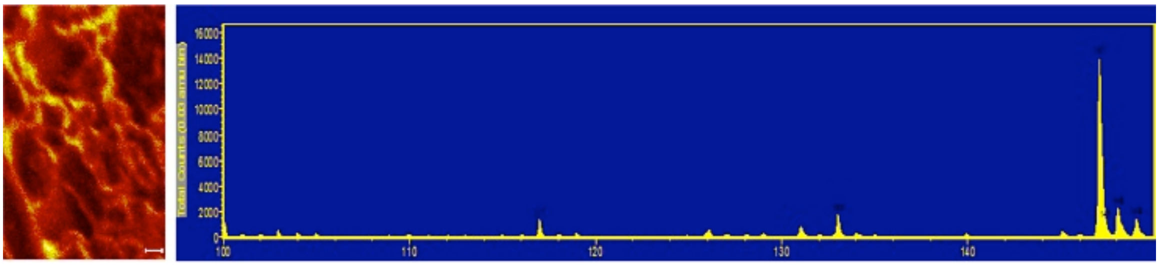


Figure 3: SIMS MSI of mCD8-Nd146 on ITO (positive mode): No signal of Nd146 could be detected if the antibody coupled tot he Nd146 was directly spotted to the ITO slide. Again, a very strong signal could be detected at m/z=147 for the likely PDMS contamination.

Although the initial experiments were not successful we are convinced that the CyToF reagents can be a very good tool for the MSI field. Therefore, we plan to perform a couple more experiments to test their suitability. First of all, we will test the pre-labeled antibody by LC-ESI MS/MS.

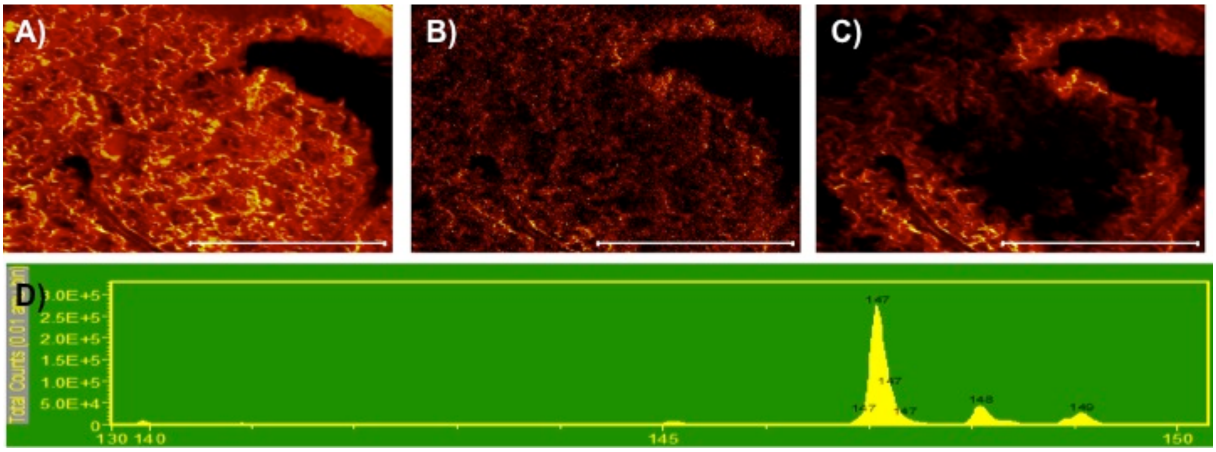


Figure 4: Dynamic SIMS of anti-CD8-Nd146 (positive mode): A) Total ion image of the antibody spot **B)** Ion image of m/z=146 **C)** Ion image of m/z=147 **D)** zoomed spectra (130-160m/z) of image A). Most abundant peak at 147m/z being most likely PDMS. Even dynamic SIMS does not show any peak at 146m/z

As described in literature there are about 30-50 polymers attached to each antibody. To each polymer about 30 Ln are coupled to each polymer (*84*). With this information we can estimate how strong the signal from the Nd should be. Additionally, we will couple Nd to the polymer and measure the Nd before and after the coupling. This should give us some indication of how strong the chelating polymer affects the ionization properties of the metal. We then label antibodies with this Nd coupled polymer and measure the signal again. If we can detect the Ln on the self-coupled antibody we will try to genetically boost the expression of the target protein to increase the possible binding partners of our antibody to proof the usefulness of Ln couples antibodies in SIMS MSI.

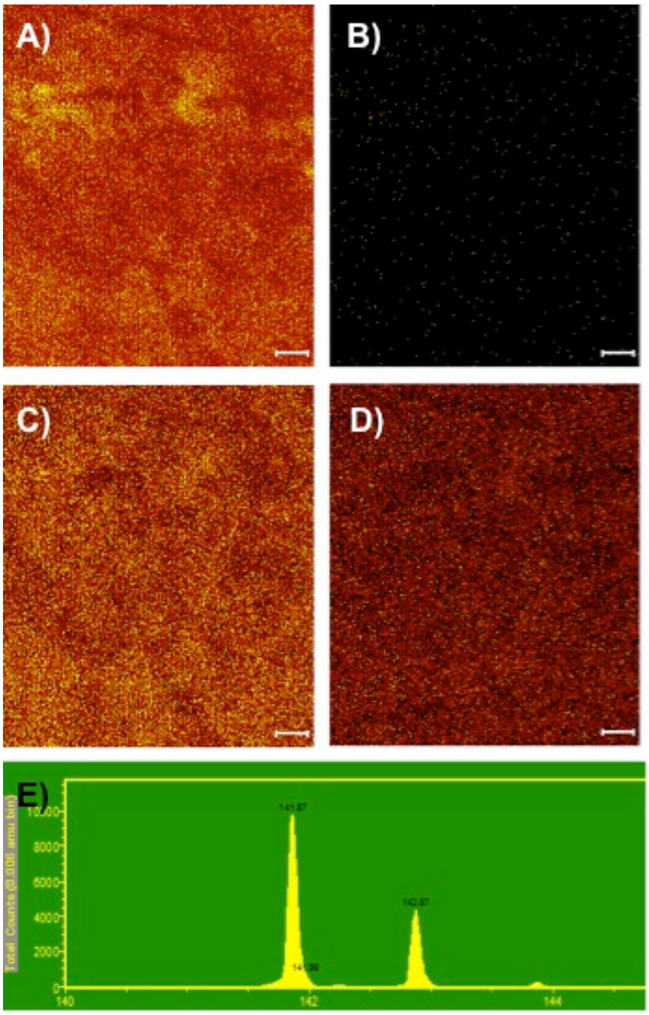


Figure 5: SIMS MS of NdCl (positive mode): A) Total ion image of the NdCl spot **B)** selected ion image of m/z=115 **C)** Ion image of m/z=142 representing the Nd142 used in this experimental series **D)** Ion image of 143m/z of either the NdH^+ or the Nd143 isotope **E).** Zoomed in spectra of A) (140-150m/z) with the most abundant peak at m/z 146 and a peak with about half the intensity at 147m/z

References

1. L. Wolpert, *Principles of Development* (Oxford University Press, (1998).

2. P. A. Lawrence, *The Making of a Fly: The Genetics of Animal Design* (Wiley-Blackwell, (1992).

3. M. Ashburner, Drosophila. A laboratory handbook., (1989).

4. R. J. Greenspan, *Fly Pushing: The Theory and Practice of* Drosophila *Genetics, Second Edition* (Cold Spring Harbor Laboratory Press, (2004).

5. C. Potter, Mechanisms of size control, *Curr. Opin. Genet. Dev.* **11**, 279–286 (2001).

6. C. Dahmann, A. C. Oates, M. Brand, Boundary formation and maintenance in tissue development., *Nat. Rev. Genet.* **12**, 43–55 (2011).

7. A. Garcia-Bellido, P. Santamaria, Developmental analysis of the wing disc in the mutant engrailed of Drosophila melanogaster, *Genetics* **72**, 87–104 (1972).

8. T. Tabata, Genetics of morphogen gradients., *Nat. Rev. Genet.* **2**, 620–30 (2001).

9. D. Nellen, R. Burke, G. Struhl, K. Basler, Direct and long-range action of a DPP morphogen gradient., *Cell* **85**, 357–68 (1996).

10. I. Guillén *et al.*, The function of engrailed and the specification of Drosophila wing pattern., *Development* **121**, 3447–56 (1995).

11. S. Schilling *et al.*, S. Shvartsman, Ed. Cell-sorting at the A/P boundary in the Drosophila wing primordium: a computational model to consolidate observed non-local effects of Hh signaling., *PLoS Comput. Biol.* **7**, (2011).

12. C. Dahmann, K. Basler, Compartment boundaries: at the edge of development, *Trends Genet.* **15**, 320–326 (1999).

13. M. Aliee *et al.*, Physical Mechanisms Shaping the Drosophila Dorsoventral Compartment Boundary, *Curr. Biol.* (2012).

14. S. S. Blair, D. L. Brower, J. B. Thomas, M. Zavortink, The role of apterous in the control of dorsoventral compartmentalization and PS integrin gene expression in the developing wing of Drosophila, *Dev.* **120** , 1805–1815 (1994).

15. F. J. Diaz-Benjumea, S. M. Cohen, Serrate signals through Notch to establish a Wingless-dependent organizer at the dorsal/ventral compartment boundary of the Drosophila wing., *Development* **121**, 4215–25 (1995).

16. D. Doherty, G. Feger, S. Younger-Shepherd, L. Y. Jan, Y. N. Jan, Delta is a ventral to dorsal signal complementary to Serrate, another Notch ligand, in Drosophila wing formation., *Genes Dev.* **10**, 421–34 (1996).

17. M. Zecca, K. Basler, G. Struhl, Direct and long-range action of a wingless morphogen gradient., *Cell* **87**, 833–44 (1996).

18. C. J. Neumann, S. M. Cohen, Long-range action of Wingless organizes the dorsal-ventral axis of the Drosophila wing., *Development* **124**, 871–80 (1997).

19. M. Zecca, G. Struhl, Subdivision of the Drosophila wing imaginal disc by EGFR-mediated signaling, *Development* **129**, 1357–1368 (2002).

20. U. Tepass, D. Godt, R. Winklbauer, Cell sorting in animal development: signalling and adhesive mechanisms in the formation of tissue boundaries, *Curr Opin Genet Dev* **12**, 572–582 (2002).

21. R. Diez del Corral, P. Aroca, J. L. G mez-Skarmeta, F. Cavodeassi, J. Modolell, The Iroquois homeodomain proteins are required to specify body wall identity in Drosophila., *Genes Dev.* **13**, 1754–61 (1999).

22. E. Villa-Cuesta, E. Gonzalez-Perez, J. Modolell, Apposition of iroquois expressing and non-expressing cells leads to cell sorting and fold formation in the Drosophila imaginal wing disc, *BMC Dev. Biol.* **7**, 106 (2007).

23. E. Villa-Cuesta, J. Modolell, Mutual repression between msh and Iro-C is an essential component of the boundary between body wall and wing in Drosophila, *Development* **132**, 4087–4183 (2005).

24. M. S. Steinberg, Reconstruction of Tissues by Dissociated Cells, *Science.,* **141**, 401–408 (1963).

25. A. Nose, A. Nagafuchi, M. Takeichi, Expressed recombinant cadherins mediate cell sorting in model systems., *Cell* **54**, 993–1001 (1988).

26. I. Heyman, A. Kent, A. Lumsden, Cellular morphology and extracellular space at rhombomere boundaries in the chick embryo hindbrain., *Dev. Dyn.* **198**, 241–53 (1993).

27. D. A. O'Brochta, P. J. Bryant, A zone of non-proliferating cells at a lineage restriction boundary in Drosophila., *Nature* **313**, 138–41.

28. S. S. Blair, Mechanisms of compartment formation: evidence that non-proliferating cells do not play a critical role in defining the D/V lineage restriction in the developing wing of Drosophila., *Development* **119**, 339–51 (1993).

29. K. P. Landsberg *et al.*, Increased cell bond tension governs cell sorting at the Drosophila anteroposterior compartment boundary., *Curr. Biol.* **19**, 1950–5 (2009).

30. G. Siuzdak, *The Expanding Role of Mass Spectrometry in Biotechnology, Second Edition* (MCC Press) (2006).

31. J. B. Fenn, M. Mann, C. K. Meng, S. F. Wong, C. M. Whitehouse, Electrospray ionization for mass spectrometry of large biomolecules., *Science* **246**, 64–71 (1989).

32. M. Karas, F. Hillenkamp, Laser desorption ionization of proteins with molecular masses exceeding 10,000 daltons., *Anal. Chem.* **60**, 2299–301 (1988).

33. B. Domon, R. Aebersold, Mass spectrometry and protein analysis., *Science* **312**, 212–7 (2006).

34. J. H. Jungmann, R. M. A. Heeren, Detection systems for mass spectrometry imaging: a perspective on novel developments with a focus on active pixel detectors., *Rapid Commun. Mass Spectrom.* **27**, 1–23 (2013).

35. K. Chughtai, R. Heeren, Mass spectrometric imaging for biomedical tissue analysis, *Chem. Rev.* **110**, 3237–3277 (2010).

36. D. F. Smith, E. W. Robinson, A. V Tolmachev, R. M. A. Heeren, L. Paša-Tolić, C60 secondary ion Fourier transform ion cyclotron resonance mass spectrometry., *Anal. Chem.* **83**, 9552–6 (2011).

37. B. Cillero-Pastor, G. Eijkel, A. Kiss, F. Blanco Garcia, R. Heeren, A time of flight secondary ion mass spectrometry based molecular distribution distinguishes healthy and osteoarthritic human cartilage, *Anal. Chem.* (2012).

38. D. Debois, M.-P. Bralet, F. Le Naour, A. Brunelle, O. Laprevote, In situ lipidomic analysis of nonalcoholic fatty liver by cluster TOF-SIMS imaging, *Anal. Chem.* **81**, 2823–2831 (2009).

39. D. Cornett, M. Reyzer, P. Chaurand, R. Caprioli, MALDI imaging mass spectrometry: molecular snapshots of biochemical systems, *Nat. Methods* **4**, 828–833 (2007).

40. E. Amstalden van Hove, D. Smith, R. Heeren, A concise review of mass spectrometry imaging, *J. Chromatogr. A* **1217**, 3946–4000 (2010).

41. L. McDonnell, R. Heeren, Imaging mass spectrometry, *Mass Spectrom. Rev.* **26**, 606–643.

42. J. Clerc, C. Fourré, P. Fragu, SIMS microscopy: methodology, problems and perspectives in mapping drugs and nuclear medicine compounds., *Cell Biol. Int.* **21**, 619–33 (1997).

43. B. Garrison, J. Vickerman, D. Briggs, ToF-SIMS: surface analysis by mass spectrometry, *Vickerman, JC, Briggs, D., Eds* , 223–257 (2001).

44. J. C. Vickerman, D. Briggs, *ToF-SIMS : Surface Analysis by Mass Spectrometry* IM Publications LLP, (2001).

45. A. Benninghoven, F. G. Rudenauer, H. W. Werner, Secondary ion mass spectrometry: basic concepts, instrumental aspects, applications and trends, (1987)

46. R. M. A. Heeren *et al.*, Why don't biologists use SIMS?, *Appl. Surf. Sci.* **252**, 6827–6835 (2006).

47. A. F. M. Altelaar *et al.*, Gold-enhanced biomolecular surface imaging of cells and tissue by SIMS and MALDI mass spectrometry., *Anal. Chem.* **78**, 734–42 (2006).

48. A. Delcorte, P. Bertrand, Interest of silver and gold metallization for molecular SIMS and SIMS imaging, *Appl. Surf. Sci.* **231**, 250–255 (2004).

49. R. J. Cotter, Time-of-flight mass spectrometry: an increasing role in the life sciences., *Biomed. Environ. Mass Spectrom.* **18**, 513–32 (1989).

50. I. M. Taban *et al.*, Imaging of peptides in the rat brain using MALDI-FTICR mass spectrometry., *J. Am. Soc. Mass Spectrom.* **18**, 145–51 (2007).

51. E. Jones, S.-O. Deininger, P. Hogendoorn, A. Deelder, L. McDonnell, Imaging Mass Spectrometry Statistical Analysis, *J. Proteomics* (2012).

52. I. Klinkert *et al.*, Tools and strategies for visualization of large image data sets in high-resolution imaging mass spectrometry., *Rev. Sci. Instrum.* **78**, 053716 (2007).

53. G. McCombie, D. Staab, M. Stoeckli, R. Knochenmuss, Spatial and spectral correlations in MALDI mass spectrometry images by clustering and multivariate analysis, *Anal. Chem.* **77**, 6118–6124 (2005).

54. L. Klerk, A. Broersen, I. Fletcher, R. van, Extended data analysis strategies for high resolution imaging MS: New methods to deal with extremely large image hyperspectral datasets, *Int. J. ,* (2007).

55. B. T. Wickes, Y. Kim, D. G. Castner, Denoising and multivariate analysis of time-of-flight SIMS images, *Surf. Interface Anal.* **35**, 640–648 (2003).

56. B. Tyler, G. Rayal, D. Castner, Multivariate analysis strategies for processing ToF-SIMS images of biomaterials, *Biomaterials* **28**, 2412–2423 (2007).

57. R. Goodwin, Sample preparation for mass spectrometry imaging: small mistakes can lead to big consequences, *J. Proteomics* **75**, 4893–4911 (2012).

58. M. Stoeckli, D. Staab, A. Schweitzer, Compound and metabolite distribution measured by MALDI mass spectrometric imaging in whole-body tissue sections, *Int. J. Mass Spectrom.* **260**, 195–202 (2007).

59. L. MacAleese, J. Stauber, R. M. A. Heeren, Perspectives for imaging mass spectrometry in the proteomics landscape., *Proteomics* **9**, 819–34 (2009).

60. B. K. Kaletaş *et al.*, Sample preparation issues for tissue imaging by imaging MS., *Proteomics* **9**, 2622–33 (2009).

61. G. Morata, P. A. Lawrence, Control of compartment development by the engrailed gene in Drosophila., *Nature* **255**, 614–7 (1975).

62. M. Ng, F. J. Diazbenjumea, S. M. Cohen, Nubbin Encodes a Pou-Domain Protein Required for Proximal-Distal Patterning in the Drosophila Wing, *Development* **121**, 589–599 (1995).

63. D. Touboul *et al.*, MALDI-TOF and cluster-TOF-SIMS imaging of Fabry disease biomarkers, *Int. J. Mass Spectrom.* **260**, 158–165 (2007).

64. D. Touboul, A. Brunelle, O. Lapr√©vote, Mass spectrometry imaging: Towards a lipid microscope?, *Biochimie* **93**, 113–119 (2011).

65. M. Bot *et al.*, Lysophosphatidic acid triggers mast cell-driven atherosclerotic plaque destabilization by increasing vascular inflammation., *J. Lipid Res.* **54**, 1265–74 (2013).

66. L. A. McDonnell, R. M. A. Heeren, Imaging mass spectrometry., *Mass Spectrom. Rev.* **26**, 606–43 (2007).

67. S. Rangarajan, B. J. Tyler, Topography in secondary ion mass spectroscopy images, *J. Vac. Sci. Technol. A Vacuum, Surfaces, Film.* **24**, 1730–1736 (2006).

68. B. Tyler, G. Rayal, D. Castner, Multivariate analysis strategies for processing ToF-SIMS images of biomaterials., *Biomaterials* **28**, 2412–2423 (2007).

69. L. A. McDonnell *et al.*, Using matrix peaks to map topography: Increased mass resolution and enhanced sensitivity in chemical imaging, *Anal. Chem.* **75**, 4373–4381 (2003).

70. J. P. R. Day, G. Rago, K. F. Domke, K. P. Velikov, M. Bonn, Label-free imaging of lipophilic bioactive molecules during lipid digestion by multiplex coherent anti-Stokes Raman scattering microspectroscopy., *J. Am. Chem. Soc.* **132**, 8433–9 (2010).

71. A. M. Hicks, C. J. DeLong, M. J. Thomas, M. Samuel, Z. Cui, Unique molecular signatures of glycerophospholipid species in different rat tissues analyzed by tandem mass spectrometry., *Biochim. Biophys. Acta* **1761**, 1022–9 (2006).

72. Y. Kalegowda, S. Harmer, Chemometric and Multivariate Statistical Analysis of Time-of-Flight Secondary Ion Mass Spectrometry Spectra from Complex Cu-Fe Sulfides, *Anal. Chem.* (2012), doi:10.1021/ac202971y.

73. D. Graham, D. Castner, Multivariate analysis of ToF-SIMS data from multicomponent systems: the why, when, and how, *Biointerphases* **7**, 49 (2012).

74. C. Barnes *et al.*, Identifying Individual Cell Types in Heterogeneous Cultures Using SIMS Imaging with C60 Etching and Multivariate Analysis, *Anal. Chem.* (2011), doi:10.1021/ac201179t.

75. F. Bonnier, H. Byrne, Understanding the molecular information contained in principal component analysis of vibrational spectra of biological systems, *Analyst* (2011).

76. M. Passarelli, N. Winograd, Lipid imaging with time-of-flight secondary ion mass spectrometry (ToF-SIMS), *Biochim. Biophys. Acta* **1811**, 976–1066 (2011).

77. X. Guan *et al.*, Biochemical Membrane Lipidomics during Drosophila Development, *Dev. Cell* **24**, 98–111 (2013).

78. M. Carvalho *et al.*, Effects of diet and development on the Drosophila lipidome, *Mol. Syst. Biol.* **8**, 600 (2012).

79. X. Gao *et al.*, A reversed-phase capillary ultra-performance liquid chromatography-mass spectrometry (UPLC-MS) method for comprehensive top-down/bottom-up lipid profiling, *Anal. Bioanal. Chem.* **402**, 2923–2933 (2012).

80. D. F. Smith *et al.*, High mass accuracy and high mass resolving power FT-ICR secondary ion mass spectrometry for biological tissue imaging., *Anal. Bioanal. Chem.* **405**, 6069–76 (2013).

81. D. F. Smith *et al.*, Advanced mass calibration and visualization for FT-ICR mass spectrometry imaging., *J. Am. Soc. Mass Spectrom.* **23**, 1865–72 (2012).

82. D. T. Scott *et al.*, Flow cytometer with mass spectrometer detection for massively multiplexed single-cell biomarker assay, *Pure Appl. Chem.* **80** (2008).

83. D. Bandura *et al.*, Mass cytometry: technique for real time single cell multitarget immunoassay based on inductively coupled plasma time-of-flight mass spectrometry, *Anal. Chem.* **81**, 6813–6822 (2009).

84. X. Lou *et al.*, Polymer-based elemental tags for sensitive bioassays, *Angew. Chem. Int. Ed. Engl.* **46**, 6111–6114 (2007).

Lebenslauf

Name	Marty
Vorname	Florian
Geburtsdatum	04.01.1983
Adresse und Wohnort	Flurstrasse 4A, 8048 Zürich
Telefonnummer	078 854 60 74
Heimatort	Guttet - Feschel (VS)
Bürgerort	Gampel-Bratsch (VS)
Zivilstand	ledig
Konfession	-

Schulische Ausbildung

1990 – 1996	Primarschule Gampel
1996 – 1998	Sekundarschule OS Gampel
1998 – 2003	Gymnasium Brig
2003 – 2004	ETH Zürich, Grundstudium Biologie
2005 – 2008	Grundstudium und Bachelor in Biologie, Universität Zürich
2008 – 2009	Master of Science in Biology/Developmental Biology, Universität Zürich

Berufliche Tätigkeit

01.09.2009 – aktuell	Doktorarbeit (PhD) in Developmental Biology, Universität Zürich im Labor von Prof. Konrad Basler

Acknowledgments

All the work performed during the duration of my PhD thesis was only possible due to the help of several people.

First and foremost, I would like to thank Dr. Erich Brunner, Prof. Dr. Konrad Basler and Prof. Dr. Ron M.A. Heeren for providing me with such an interesting project to work on and for their continuous support, input and discussion during my PhD.

Special thanks belong to Dr. Luke MacAleese who was kind and patient enough to introduce me into the field of imaging mass spectrometry.

Further, Dr. Don F. Smith was of tremendous help during the later stage of the project as he was always willing to discuss problems on and off topic. Additionally, he performed the first high mass resolution FT data on the wing imaginal disc at PNNL.

For the part on CARS spectroscopy I thank the lab of Prof. Dr. Mischa Bonn for their fruitful collaboration. Special thanks to Dr. Gianluca Rago who performed all the CARS experiments that are part of this work.

The analysis and data interpretation of the lipid catalogue generated was only possible because of the help of Dr. Xiaoli Gao. Many thanks for helping out on that part of my project.

Additional thanks to Prof. Dr. Reinhard Furrer and Florian Gerber for their effort during our collaborative work together with Gert Eijkel.

I would also like to say thank you to Dr. Markus Stoeckli for his input during the committee meetings. Additionally, he gave me the chance to work together with Dieter Staab in the laboratories at Novartis, Basel.

Further I would like to thank the Basler and Heeren lab for their good working atmosphere they provided.

Special thanks belongs to Franklin E. Leach, Christopher S. Anderton and Lilijana Pasa-Tolic for their work on the high mass resolution C60 FT ICR data.

I would also like to thank all the secretaries at the Institute of Molecular Life Sciences (UZH), FOM/AMOLF Institute and EMSL/PNNL. Their help was of fundamental importance to work at all these different places.

Special thanks also to all my friends who were always there to support me, were patient with me and helped me along this journey.

To Dr. Kerstin Euler I am thankful for reading carefully through the draft of this thesis and giving me a lot of helpful advice, critics, comments and input.

Last but not least a big thank you to my family. My brother Patrick, my sister Katja and my parents Miranda and Erich were always supporting me.

As many people have helped me during the last couple of years I am certain that I forgot some of you to mention here. For all those that provided help, discussion, critics, input, feedback, reagents, flies or a simple beer to forget all the troubles and problems, I would simply like to say:

Thank You All!!!

- What is now proved was once only imagined. -

William Blake (1757 – 1827)